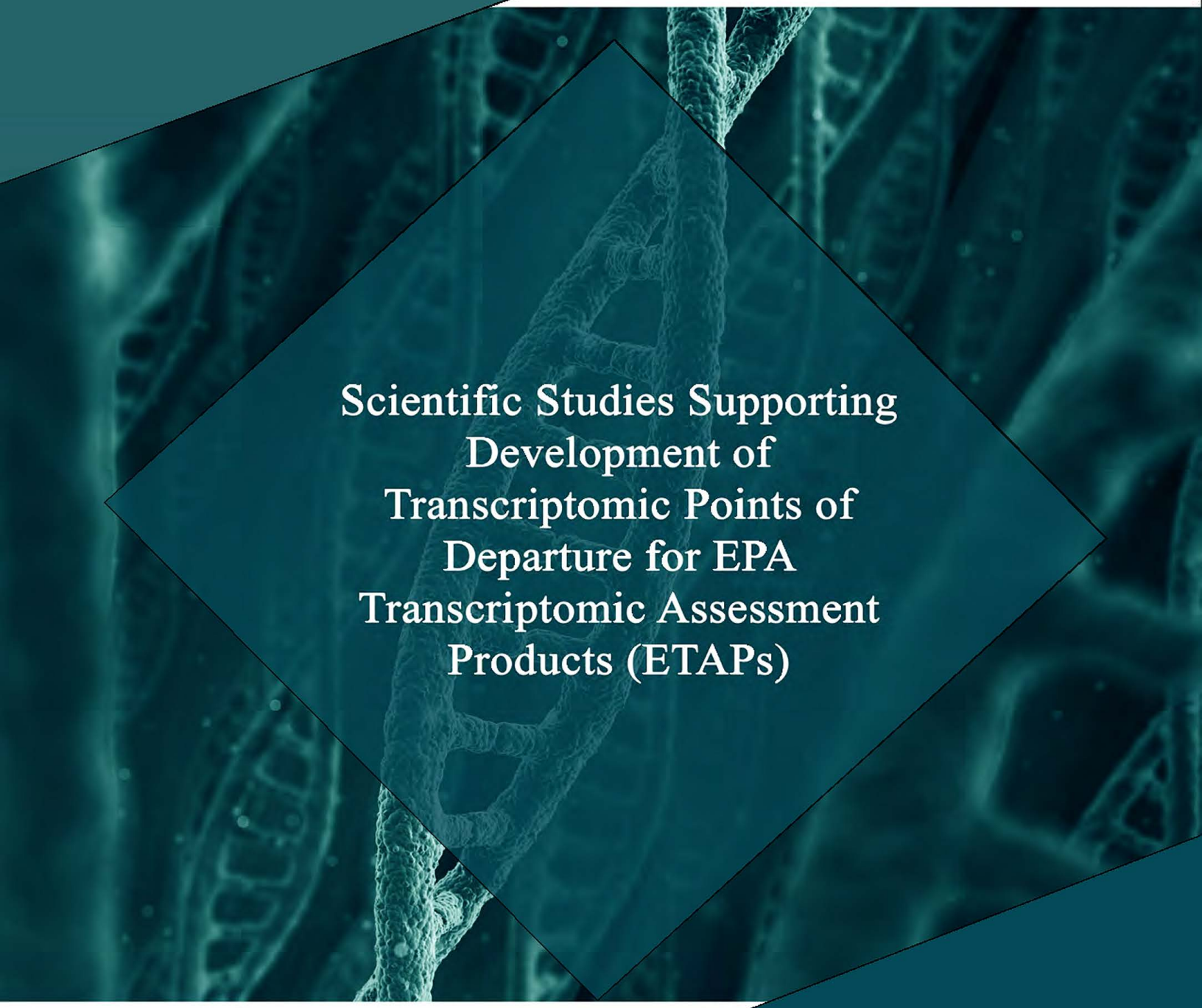
United States Environmental Protection Agency

Final Report



Scientific Studies Supporting Development of Transcriptomic Points of Departure for EPA Transcriptomic Assessment Products (ETAPs)

EPA/600/X-23/084 | March 2024 | www.epa.gov/research
https://doi.org/10.23645/epacomptox.25365550

Office of Research and Development

EPA/600/X-23/084

Scientific Studies Supporting Development of Transcriptomic Points of Departure for EPA Transcriptomic Assessment Products (ETAPs)

March 2024

Center for Computational Toxicology and Exposure (CCTE) &
Center for Public Health and Environmental Assessment (CPHEA)
Office of Research and Development
U.S. Environmental Protection Agency

DISCLAIMER

This document has been reviewed in accordance with the U.S. Environmental Protection Agency's (EPA) peer and administrative review policies and approved for publication. Peer review was performed by the EPA Board of Scientific Counselors and was subject to public comment. Any mention of trade names, products, or services does not imply an endorsement by the U.S. government or the EPA. EPA does not endorse any commercial products, services, or enterprises.

TABLE OF CONTENTS

ABBREVIATIONS

ANOVA	Analysis of Variance
AIC	Akaike Information Criterion
ATSDR	Agency for Toxic Substances and Disease Registry
BCTD	Biomolecular and Computational Toxicology Division
BMD	Benchmark Dose
BMD(L)	Refers to the BMD and/or BMDL
BMDL	Benchmark Dose Lower Confidence Bound
BMDS	Benchmark Dose Modeling Software
BMDU	Benchmark Dose Upper Confidence Bound
BMR	Benchmark Response
BOSC	EPA Board of Scientific Counselors
CASRN	Chemical Abstracts Service Registry Number
CCCB	Computational Chemistry and Cheminformatics Branch
CCED	Chemical Characterization and Exposure Division
CCTE	Center for Computational Toxicology and Exposure
CDx	Companion Diagnostics
CPAD	Chemical and Pollutant Assessment Division
CPHEA	Center for Public Health and Environmental Assessment
CPM	Counts per Million
CTBB	Computational Toxicology and Bioinformatics Branch
DNTP	Division of the National Toxicology Program of the National Institute of Environmental Health Sciences
DTT	Division of Translational Toxicology of the National Institute of Environmental Health Sciences, formerly known as the Division of the National Toxicology Program (DNTP)
DWS	Drinking Water Standards
EPA	U.S. Environmental Protection Agency
ECHA	European Chemicals Agency
ENBS	Expected Net Benefit of Sampling
ETAP	EPA Transcriptomic Assessment Product
ETTB	Experimental Toxicokinetics and Toxicodynamics Branch
FC	Fold Change
FDA	U.S. Food and Drug Administration
FDR	False Discovery Rate
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
GO	Gene Ontology
ID	Identifier
IRIS	EPA Integrated Risk Information System
KOW	n-Octanol/Water Partition Coefficient
LOAEL	Lowest Observable Adverse Effect Level
MAQC	The MicroArray Quality Consortium
MRL	Minimum Risk Level
mRNA	Messenger Ribonucleic Acid (RNA)
MSD	Mean Squared Difference
MAD	Median Absolute Deviation
NAM	New Approach Methodology
NASEM	U.S. National Academies for Science, Engineering, and Medicine
NGS	Next-Generation Sequencing
NIEHS	National Institute of Environmental Health Sciences

NIH	National Institutes of Health
NOAEL	No Observable Adverse Effect Level
NTP	National Toxicology Program
OECD	Organization for Economic Cooperation and Development
OPP	EPA Office of Pesticide Program
ORD	Office of Research and Development
OW	EPA Office of Water
PBPK	Physiologically Based Pharmacokinetic
PC/PCA	Principal Component / Principal Component Analysis
PFAS	Per- and Polyfluoralkyl Substances
POD	Point-of-Departure
PPARα	Peroxisome Proliferator Activated Receptor α
PPRTV	Provisional Peer Reviewed Toxicity Value
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals
RfD	Reference Dose
RMSD	Root- Mean-Square Difference
RMSE	Root- Mean-Square Error
RNA-Seq	Ribonucleic Acid (RNA) Sequencing
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SD	Standard Deviation
SEQC	Sequencing Quality Control
TEAB	Toxic Effects Assessment Branch
TG-GATES	Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System
THHA	Traditional Human Health Assessment
TSCA	Toxic Substances Control Act
TRV	Transcriptomic Reference Value
UF	Uncertainty Factor
VOI	Value of Information

AUTHORS | CONTRIBUTORS

EPA Office of Research and Development (ORD) Authors

Dan Chang	EPA/ORD/CCTE
John Cowden	EPA/ORD/CCTE
Sarah Davidson-Fritz	EPA/ORD/CCTE
Jeffry Dean	EPA/ORD/CPHEA
Mike Devito	EPA/ORD/CCTE
Logan Everett	EPA/ORD/CCTE
Alison Harrill	EPA/ORD/CCTE
Susan Hester	EPA/ORD/CCTE (retired)
Michael Hughes	EPA/ORD/CCTE
Jason Lambert	EPA/ORD/CCTE
Lucina Lizarraga	EPA/ORD/CPHEA
Roman Mezencev	EPA/ORD/CPHEA
Grace Patlewicz	EPA/ORD/CCTE
Kristina Thayer	EPA/ORD/CPHEA
Russell Thomas	EPA/ORD/CCTE
Kelsey Vitense	EPA/ORD/CCTE
Leah Wehmas	EPA/ORD/CCTE

NIEHS Division of Translational Toxicology (DTT) Authors

Scott Auerbach	NTP/DTT/NIEHS/NIH
Warren Casey	NTP/DTT/NIEHS/NIH

EPA ORD Executive Direction

Russell Thomas	CCTE Center Director
Reeder Sams	CCTE Deputy Center Director
Jason Lambert	CCTE Senior Science Advisor
Alison Harrill	CCTE Associate Director
Mike Devito	CCTE/CCED Division Director
Sid Hunter	CCTE/BCTD Division Director
Michael Hughes	CCTE/CCED/ETTB Branch Chief
Dan Chang	CCTE/CCED/CCCB Branch Chief

John Cowden	CCTE/BCTD/CTBB Branch Chief
Wayne Cascio	CPHEA Center Director
Kay Holt	CPHEA Deputy Center Director
Samantha Jones	CPHEA Associate Director
Kris Thayer	CPHEA/CPAD Division Director
Ravi Subramaniam	CPHEA/CPAD/TEABD Branch Chief
Glenn Rice	CPHEA/CPAD/TEABC Branch Chief

EPA and DTT Contributors and Reviewers

John Bucher	NTP/DTT/NIEHS/NIH
Timothy Buckley	EPA/ORD/CCTE
Peter Egeghy	EPA/ORD/CCTE
Katie Paul-Friedman	EPA/ORD/CCTE
Joshua Harrill	EPA/ORD/CCTE
Sid Hunter	EPA/ORD/CCTE
Kristin Isaacs	EPA/ORD/CCTE
Richard Judson	EPA/ORD/CCTE
Jonathan Kaiser	EPA/ORD/CPHEA
Scott Masten	NTP/DTT/NIEHS/NIH
Fred Parham	NTP/DTT/NIEHS/NIH
Grace Patlewicz	EPA/ORD/CCTE
Dan Petersen	EPA/ORD/CPHEA
Allison Phillips	EPA/ORD/CPHEA
Jon Sobus	EPA/ORD/CCTE
Daniel Villeneuve	EPA/ORD/CCTE
Paul White	EPA/ORD/CPHEA
Antony Williams	EPA/ORD/CCTE
George Woodall	EPA/ORD/CPHEA
Jay Zhao	EPA/ORD/CPHEA

1.EXECUTIVE SUMMARY

Current estimates of the size of worldwide and domestic chemical inventories approach hundreds of thousands of chemicals, with increasing trends in future chemical production and release. Relatively few of the chemicals in commerce, or those found in the environment, various waste streams, and the human body, have traditional toxicity data and fewer have human health assessments. Given historical, current, and future trends in chemical production and the lack of toxicity testing data available to inform human health assessments, the U.S. Environmental Protection Agency (EPA) is frequently faced with making decisions with limited or no data when evaluating potential human health risks.

Transcriptomics is the large-scale measurement of gene expression changes, and its application to toxicology enables broad characterization of the biological processes and pathways that may be impacted following exposure to a chemical. The technology and analysis methods for characterizing transcriptomic responses have matured and moved beyond the research laboratory into regulatory application. A literature review was conducted to evaluate the potential for using transcriptomic points of departure (PODs) from short-term *in vivo* studies in rodents to predict apical PODs from traditional *in vivo* toxicity studies. The literature survey included over 140 chemicals with diverse properties tested in 33 independent studies with varying experimental designs. The results of the literature survey demonstrated that transcriptomic benchmark dose (BMD) and benchmark dose lower confidence bound (BMDL) values, when integrated at a gene set level, were consistently concordant with BMD and BMDL values for apical responses in traditional subchronic and chronic rodent toxicity studies. The transcriptomic and apical dose concordance was robust across different exposure durations, exposure routes, species, sex, target tissues, physicochemical properties, toxicokinetic half-lives, and technology platforms. For the 40 chemicals with reported chronic rodent bioassay results, the Pearson's correlation coefficient was 0.820 with a $\log_{10}$ root-mean-square difference (RMSD) of 0.593 ($\log_{10}$ mg/kg-day) and a median absolute ratio of 2.4 ± 1.0 (Median Absolute Deviation; MAD). The RMSD value is similar to the range of inter-study standard deviation estimates for the lowest observable adverse effect levels (LOAELs) for systemic toxicity in repeated dose studies, approximated as residual root-mean-square error (RMSE) in log10-mg/kg-day units [0.45-0.56; (Pham et al. 2020)]. The results suggest that the error associated with the concordance between the transcriptomic BMD values versus non-cancer and cancer apical BMD values is approximately equivalent to the inter-study variability in the repeated dose toxicity study itself.

Building on the results of the literature survey, the EPA Office of Research and Development (ORD) evaluated a process to derive transcriptomic PODs based on recommendations in the peer-reviewed National Toxicology Program (NTP) Approach to Genomic Dose Response Modeling report (NTP 2018). The methodology utilizes a 5-day, repeated dose *in vivo* study in male and female rats

with an extended dose response series. The transcriptomic dose response modeling follows a stepwise process that utilizes BMD modeling approaches that are commonly employed in chemical risk assessment. In the evaluation process, a comprehensive series of analyses was performed to identify and support the choices and parameters used in each step of the transcriptomic dose response modeling process to promote detection of true signal, maximize reproducibility, and minimize false signal. A combination of parameters was identified that resulted in a Pearson correlation coefficient and $\log_{10}$ RMSD for the transcriptomic and chronic apical BMD values of 0.910 and 0.567, respectively. The median absolute ratio of the transcriptomic BMD and chronic apical BMD values was 3.2 $\pm$ 1.9 (MAD). The inter-study transcriptomic BMD standard deviation for a subset of chemicals that were independently replicated was 0.242 ($\log_{10}$ mg/kg-day). The overall estimated family-wise error rate for identifying a gene set level BMD was 0.006.

To provide context for the transcriptomic and apical BMD concordance, a statistical analysis was conducted to derive a lower bound of the expected mean squared difference (MSD) given inter-study variances in both the transcriptomic and apical responses. The results showed that the MSD of the transcriptomic and apical BMD concordance for the top performing combination of parameters (ranging from 0.285 - 0.387, depending on chemical replicates used) falls within the expected range of 0.267 - 0.617 ($\log_{10}$ mg/kg-day)^2 when considering inter-study variances. The results of the analysis suggest that the error associated with the concordance between the transcriptomic BMD values versus non-cancer and cancer apical BMD values is approximately equivalent to the combined inter-study variability associated with the 5-day transcriptomic study and the two-year rodent bioassay.

The overall conclusions from the literature survey, evaluation of the transcriptomic dose response analysis methods, and the statistical comparison of the concordance with inter-study variances support the use of transcriptomic PODs from 5-day, repeated dose *in vivo* rodent studies in quantitative human health assessments. The application is supported based on multiple studies demonstrating transcriptomics as a reliable method to measure changes in gene expression; extensive peer review of the study design and dose response analysis methods in the individual publications and NTP report; availability of peer-reviewed software for reproducible application across datasets; broad application of the use of these dose response analysis methods in the government, academic, and private sectors; and historical precedence of the underlying dose response modeling methods in risk assessment. The application is further supported based on the performance of the method in approximating an apical POD from two-year toxicity studies, an inter-study variability that is consistent with those estimated for repeated chronic toxicity studies, and low family-wise error rate.

2. BACKGROUND

2.1. WORLDWIDE AND DOMESTIC CHEMICAL LANDSCAPE

Modern societies are reliant upon chemicals to provide raw inputs for agriculture and manufacturing processes, with over 95% of manufactured goods and articles estimated to be reliant upon some form of industrial chemical process (OE 2019). A recent survey of chemical registries in 19 countries or regions found that greater than 350,000 chemicals and mixtures of chemicals were registered in one or more inventories (Wang et al. 2020). While large, this estimate of the combined worldwide chemical inventory is likely to be an undercount, as the inventories from which these numbers were derived do not include unintentionally produced materials (including unreacted intermediates, by-products, or degradation products), and the requirements of mass-production thresholds for registration lead to exclusion of certain manufactured chemicals from inventory lists. Domestically, the 2022 Toxic Substances Control Act (TSCA) chemical inventory contained more than 86,000 chemicals, of which approximately 42,000 are considered commercially active[1]. However, the global and domestic inventory numbers are a snapshot in time and trends in chemical production and use continue to rise (EEA 2018). The result of the historical, current, and future trends in chemical production is a substantial and increasing number of chemicals for regulatory agencies, such as the EPA, to evaluate for human health risks.

2.2. TOXICITY TESTING DATA LANDSCAPE

Understanding the human health impacts of exposure to chemicals requires testing and access to toxicity data. To evaluate the potential human health risk(s) of chemicals, assess potential impacts on the environment, and approve chemicals for certain uses, the EPA uses information from a broad range of animal studies, in conjunction with human data when available. While EPA and other federal agencies are working on the development and application of New Approach Methodologies (NAMs) to evaluate the potential human health hazards of chemicals (EPA 2021; NASEM 2007), toxicity testing has traditionally been performed in animal studies that employ varying exposure durations and that consider different health effect domains. The guideline animal studies include general acute, subchronic, and chronic repeat-dose toxicity studies as well as more targeted study designs such as those focused on neurotoxicity, reproductive toxicity, immunotoxicity, and developmental toxicity. The requirements for conducting these tests vary depending on the intended use of the chemical and specific statutes governing those uses. For example, conventional food-use

[1] US EPA TSCA Inventory: https://www.epa.gov/tsca-inventory

pesticide active ingredients require a full battery of animal toxicology studies. In contrast, no specific animal tests are required for registering commercial and industrial chemicals regulated under TSCA.

To understand the overall landscape of toxicity testing data relevant to human health, the availability of testing data can be tallied across different sets of chemicals that are representative of those that are present in our environment, waste streams, contaminants of immediate and emerging concern, human bodies, and commerce. For chemicals in the environment, the EPA Multimedia Monitoring Database[2] identifies chemicals in a diverse range of environmental and biological media (Isaacs et al. 2022). Databases of chemicals identified in biosolids[3] and produced water[4] (Danforth et al. 2020) provide information on substances identified in human waste and by-products derived from oil and gas extraction. For contaminants of immediate and emerging concern, the Organization for Economic Cooperation and Development (OECD) database of per- and polyfluoroalkyl substances (PFAS)[5] provides a comprehensive list of PFAS substances that may have been on the global market at some point in time (OECD 2018). A database of the blood exposome[6] that is filtered for chemicals on the TSCA inventory provides a list of non-pesticide exogenous chemicals that may be present in the human body (Barupal and Fiehn 2019). Finally, the TSCA active inventory provides substances that are currently active in US commerce[7]. Among chemicals in each of these sets, the maximum fraction of chemicals with any of the toxicity tests is 44% for chemicals in the blood exposome (Fig. 2-1). This suggests that more than half of chemicals identified in the blood have no toxicity testing data. Contaminants of immediate and emerging concern represented by PFAS are highly understudied, with only 1% PFAS substances having any available toxicity data. The more specialized toxicity tests, such as those assessing developmental toxicity, reproductive toxicity, immunotoxicity, neurotoxicity, and carcinogenicity, never exceed 30% among these representative sets of chemicals. For those chemicals identified in environmental media, only 43% have been subjected to any of the toxicity tests, while only 15% of chemicals in US commerce have been similarly tested. The results from the combined analyses on these representative sets highlight the limited human health toxicity information available to regulatory agencies, such as the EPA, on substances in the environment and to which humans are likely exposed.

[2] Multimedia Monitoring Database: https://comptox.epa.gov/dashboard/chemical-lists/MMDBV1
[3] Biosolids: https://comptox.epa.gov/dashboard/chemical-lists/BIOSOLIDS2022
[4] Produced water exists in subsurface formations and is brought to the surface during oil and gas production: https://comptox.epa.gov/dashboard/chemical-lists/PRODWATER
[5] Organization for Economic Cooperation and Development (OECD) global PFAS database: https://comptox.epa.gov/dashboard/chemical-lists/PFASOECD
[6] TSCA subset of the blood exposome: https://comptox.epa.gov/dashboard/chemical-lists/BLOODTSCA The TSCA subset of the blood exposome was used to represent the non-pesticide exogenous chemicals among those detected.
[7] US EPA TSCA Inventory: https://www.epa.gov/tsca-inventory

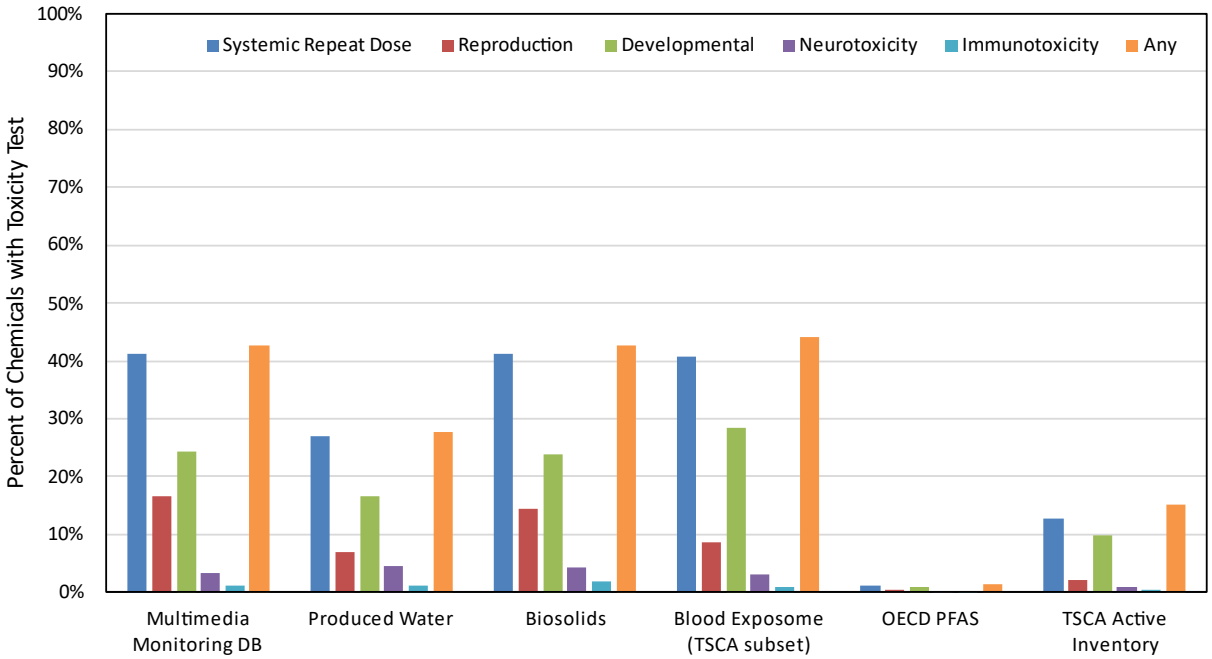


Figure 2-1. Bar graph depicting the percentage of chemicals with different toxicity tests across various representative chemical sets. The sets of chemicals were selected to represent substances found in environmental media (Multimedia Monitoring Database), different waste streams (produced water from oil and gas extraction; biosolids), the human body (TSCA subset of the blood exposome), contaminants of emerging concern (OECD PFAS list), and commerce (TSCA active inventory). The 'Systemic Repeat Dose' toxicity test includes repeat dose studies of subchronic and chronic duration. The 'Any' category is the union of unique chemicals across the various study types. The total percentages of chemicals across study types may not equal the total percentage in 'Any' given that chemicals may have multiple different studies. The percentages of chemicals with toxicity tests were calculated based on the respective studies in ToxValDB v9.4.

2.3. HUMAN HEALTH ASSESSMENT LANDSCAPE

Human health assessments are developed to identify chemical exposure levels likely to be without appreciable risk of deleterious effects during an individual's lifetime. These assessments inform a broad range of regulatory decisions, such as setting water quality standards or remediation levels at contaminated sites, restrictions on use, and standards for manufacturing, disposal, and air emissions. The lack of traditional repeated dose toxicity study data for most chemicals is one of the factors that impedes hazard identification and dose response assessment, including identification of

points of departure (PODs)[8] and derivation of corresponding reference values, which underpin human health risk assessments. Other factors that have impacted the development of human health assessments include the time and resources required to perform and review the studies, derive the resulting reference value, and publish the assessment. For example, it takes the EPA approximately 15 to 36 months to review the testing results of a single conventional agricultural pesticide following the data collection phase[9]. For industrial and commercial chemicals, the development of human health assessments can often take 4 years or more (Krewski et al. 2020), while more complex assessments can take substantially longer (NASEM 2009). These time estimates do not include performing the toxicity tests that the assessments rely upon for informing the derivation of human health reference values. The downstream consequences of the insufficient toxicity data, as well as the time and resources required to develop human health assessments, result in fewer chemicals with reference values for regulatory applications. To provide a picture of the human health assessment landscape, the availability of human health assessments from EPA and the Agency for Toxic Substances and Disease Registry (ATSDR) were tallied across the same sets of representative chemicals highlighted in Figure 2-1. These data are depicted in Figure 2-2. Across all the sets of chemicals, the maximum percentage of chemicals with a human health assessment were those identified in environmental media at 24%. For chemicals in biosolids, approximately 23% have human health assessments, while 11% of chemicals identified in human blood have assessments. Among chemicals on the TSCA active inventory, fewer than 2% of chemicals have human health assessments and 0.2% of the OECD list of PFAS have assessments. The overall conclusion from this comparison is that there remains a substantial number of chemicals in the environment and to which humans are likely exposed that do not have human health assessments. Filling this gap with traditional toxicity tests and chemical assessment workflows would require multiple decades and substantial resources to complete.

[8] In human health risk assessment practice, a point-of-departure (POD) represents the dose-response point that marks the beginning of a low-dose extrapolation. This point can be the lower bound on dose for an estimated incidence or a change in response level from a dose-response model (*e.g.*, Benchmark Dose; BMD), or a No Observable Adverse Effect Level (NOAEL) or Lowest Observable Adverse Effect Level (LOAEL) for an observed incidence, or change in level of response. For BMD values, this is typically the BMD lower confidence bound (BMDL).

[9] EPA Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Science Advisory Panel White Paper: https://www.regulations.gov/document/EPA-HQ-OPP-2011-0284-0006

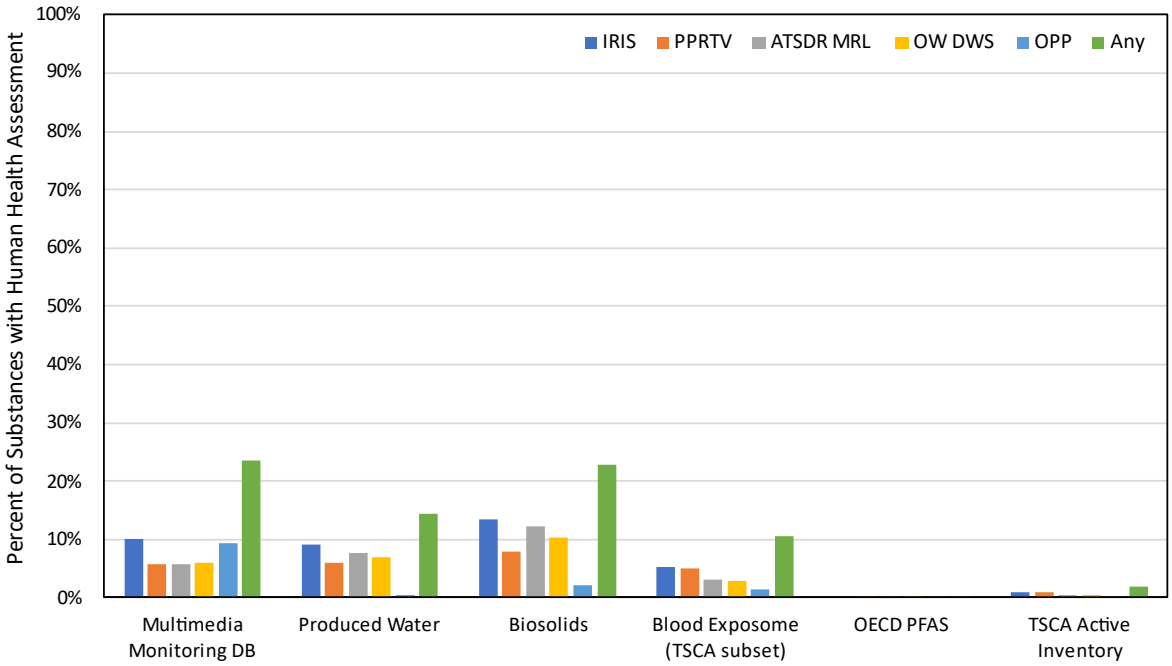


Figure 2-2. Bar graph depicting the percentage of chemicals with human health assessments from US federal agencies across various representative chemical sets. The sets of chemicals were selected to represent substances found in environmental media (Multimedia Monitoring Database), different waste streams (produced water from oil and gas extraction; biosolids), the human body (TSCA subset of the blood exposome), contaminants of emerging concern (OECD PFAS list), and the economy (TSCA active inventory). The percentages of chemicals with human health assessments were calculated based on the overlap with chemicals with non-cancer assessments in the EPA Integrated Risk Information System (IRIS) database, EPA Provisional Peer Reviewed Toxicity Values (PPRTV), Agency for Toxic Substances and Disease Registry (ATSDR) Minimum Risk Levels (MRL), EPA Office of Water Drinking Water Standards (OW DWS), and Office of Pesticide Program (OPP) risk assessments. The 'Any' category is the union of unique chemicals across the various assessment types. The total percentages of chemicals across assessment types may not equal the total percentage in 'Any' given that chemicals may have multiple different assessments (e.g., a chemical may have an IRIS and ATSDR MRL).

2.4. DEVELOPMENT OF THE EPA TRANSCRIPTOMIC ASSESSMENT PRODUCT (ETAP) TO ADDRESS THE HUMAN HEALTH ASSESSMENT GAP

Transcriptomics is the large-scale measurement of gene expression changes, and its application to toxicology enables broad characterization of the biological processes and pathways that may be impacted following exposure to a chemical. The technology has matured and has become broadly available and reproducible. Standardized methods and software applications have been developed for analyzing transcriptomic changes as a function of dose using methods traditionally applied in chemical risk assessment. In addition, a substantial body of research – detailed in Section 4 - has demonstrated that doses causing coordinated transcriptional changes are concordant with

doses causing adverse apical[10] effects in traditional toxicity studies, underscoring its potential application to regulatory decision making.

To help meet the need for toxicity testing and human health assessment of chemicals, the EPA is proposing to develop transcriptomic reference values (TRV) for use in EPA Transcriptomic Assessment Products (ETAP). The TRV is defined as an estimate of a daily oral dose to the human population that is likely to be without appreciable risk of adverse non-cancer health effects over a lifetime. The TRV is derived from a transcriptomic POD with uncertainty factors (UFs) applied to reflect limitations of the data used. The transcriptomic POD is defined as the dose at which there were no coordinated transcriptional changes that would indicate a potential toxicity of concern. The coordinated transcriptional changes used to identify the POD do not necessarily discriminate between specific hazards, adverse or adaptive effects, nor are they used to infer a mechanism or mode of action. The ETAP is intended to be applied to substances with no existing or publicly accessible repeated dose toxicity studies or human evidence suitable for use as a POD and reference value derivation. The ETAP may be updated to incorporate new data or methodologies that might impact the TRV or retired if traditional toxicity studies and an associated human health assessment are published.

The ETAP consists of three primary components: 1) initial database searches and systematic evidence map development; 2) short-term (5-day) *in vivo* transcriptomic study and POD determination; and 3) assessment development and reporting. The initial database searches and systematic evidence map are used to verify the lack of suitable *in vivo* experimental animal toxicity studies or human epidemiological studies. The transcriptomic PODs[11] obtained from the short-term *in vivo* study in rodents are used in the derivation of a TRV through application of uncertainty factors (UFs) consistent with human health risk assessment practice in the EPA (EPA 2022). The values of the individual UFs and the overall composite uncertainty factor are the same across the individual ETAP assessments due to the standardized nature of the transcriptomic studies and data analysis procedures. The results from the systematic evidence mapping, 5-day transcriptomic study, and TRV derivation are compiled and reported in a standard ETAP template. A main concept of the ETAP is that the underlying methods and data analysis procedures are more standardized and structured than typical toxicity test guidelines and human health assessment documents, and the decision context is narrowly focused on substances with no existing or publicly accessible repeated dose toxicity studies or human evidence suitable for use as a POD and reference value derivation. Due to

[10] An apical endpoint is an observable outcome in a whole organism, such as a clinical sign or pathologic state, that is indicative of a disease state that can result from exposure to a toxicant (NASEM 2007). In this document, apical endpoints also include other phenotypic responses (e.g., organ and body weight changes) that are commonly used as critical effects in chemical risk assessment.

[11] In this document, the transcriptomic POD is defined as the transcriptomic BMDL, which is calculated at the level of a gene set. Some studies in the review section may have defined the transcriptomic POD in other ways and these differences are noted, where evident.

the standardized methods, the ETAP includes a streamlined review process that is intended to facilitate the rapid development, execution, and release of the human health assessments.

2.5. VALUE OF INFORMATION (VOI) ANALYSIS OF THE POTENTIAL PUBLIC HEALTH AND ECONOMIC TRADE-OFFS OF ETAP

A key aspect of public health decision-making with respect to potentially toxic chemicals present in the environment involves the choice between making an immediate decision with currently available information versus delaying a decision until additional data are collected and analyzed. This choice is often informed by the urgency of the public health need and the costs, in terms of both time and resources, of acquiring additional relevant information that may lead to decisions with less uncertainty. The US National Academies of Sciences, Engineering, and Medicine (NASEM), in its report *Science and Decisions*: *Advancing Risk Assessment,* reflected that time is a major and rarely acknowledged factor in risk assessment and that additional studies may reduce uncertainty, but the delay can have significant impacts on society and communities who are exposed while awaiting the results (NASEM 2009). The NASEM Committee recommended the development of a value of information (VOI) analysis to provide a more objective decision framework to evaluate the potential impact of new information on a particular decision or proposed changes in risk assessment activities (NASEM 2009). Given the limited human health toxicity information available on chemicals present in our environment, waste streams, human bodies, and commerce, as well as the need to gain efficient insight into contaminants of immediate and emerging concern, it is important to evaluate the VOI that new testing and human health assessment strategies may offer.

EPA ORD has developed a VOI framework that considers the public health and economic trade-offs associated with uncertainty, duration, and cost of chemical toxicity testing (Hagiwara et al. 2022). The VOI framework was used in a case study to evaluate public health and economic value of the short-term *in vivo* transcriptomic study and ETAP compared to a two-year rodent bioassay and traditional human health assessment (THHA) process (EPA 2024a). Both the ETAP and the THHA result in the derivation of a reference value that represents the daily dose of a chemical substance for which exposure to humans would be unlikely to result in an adverse health effect following oral exposure. The ETAP and THHA process differ in the cost of the study, duration of the study and assessment process, and the degree of uncertainty around the POD, with ETAP having the inherently shorter-duration and lower cost assay, but with the trade-off of presumed greater uncertainty. The comparison was performed for various combinations of exposure conditions, health endpoints, control costs, population characteristics, and decision types. A total of 306 combinations of input parameters were developed as scenarios to capture the diverse range of characteristics and contexts for chemicals in the domain of ETAP. The scenarios and associated parameters were informed by empirical data and real-world information where available. This case study is the subject of a separate Board of Scientific Counselors (BOSC) review; however, the results of the analysis were

intended to complement the scientific review of the ETAP by providing the potential socioeconomic impacts of the draft human health assessment if implemented.

The results of the case study showed that ETAP was favored over THHA in the majority of scenarios examined and across multiple VOI metrics. The benefit-risk decision maker chooses to mitigate exposure if the reduction in health cost (or increased health benefit) outweighs the associated cost of control. When evaluated by a benefit-risk decision maker, 81% of the scenarios favored ETAP, while the remaining 19% favored neither ETAP nor THHA. The target-risk decision context takes regulatory action to mitigate exposure whenever the risk exceeds a prespecified target risk level. The target-risk decision context does not consider the cost of the exposure mitigation efforts. When evaluated by a target-risk decision maker, between 89 and 99% of the scenarios favored ETAP depending on the VOI metric evaluated with 8% of the scenarios favoring neither ETAP nor THHA. The median difference in the expected net benefit of sampling (ENBS), which considers the reduction in total costs from the additional testing and assessment activities adjusted for delay and the cost of testing, was approximately $47 billion for the benefit-risk decision context and $82 billion for the target-risk decision context[12]. Negative values for the ENBS were frequently observed for THHA and the benefit-risk decision context suggesting that the delay and costs associated with decision-making for the traditional toxicity testing and human health assessment process are greater than the eventual benefit. In contrast, the ETAP less frequently had a negative ENBS for the benefit-risk decision context, suggesting that the benefit gained by collecting toxicity information via ETAP outweighed both the delay and the cost of testing for most scenarios evaluated.

Overall, the results of the case study indicate that under the exposure scenarios and assumptions considered, the ETAP was more frequently preferred over the traditional toxicity testing and human health approach for more rapidly and cost effectively evaluating chemicals with no existing toxicity testing or human health data. The amount of time needed to conduct the toxicity testing and develop the human health assessment was particularly important when the risks were high as the delay in implementing regulatory action resulted in a large cost of delay.

2.6. BOARD OF SCIENTIFIC COUNSELORS (BOSC) SCIENTIFIC REVIEW OF ETAP DEVELOPMENT AND IMPLEMENTATION

The review of the ETAP methods by the EPA BOSC is being performed in close coordination with a separate review of the VOI case study comparing the short-term *in vivo* transcriptomic study and ETAP with the traditional chronic rodent bioassay and human health assessment process. The series of white papers are intended to be complementary and cover not only the scientific support for the proposed toxicity test and assessment product, but also the socioeconomic rationale.

[12] Total benefits over the twenty-year time horizon for which the costs were calculated. To put these numbers in perspective, as an example, preventing a chronic health condition (costing $10,000 per year) among 330,000 people (approximately 1/10th of 1% of the US population) would provide a benefit of $66B over a 20-year time horizon.

For the ETAP, the review package to support the development and implementation of the assessment consists of two main documents. The first document (this one) is focused on the scientific studies and analyses supporting development of transcriptomic PODs for ETAP. Herein, the document is organized to provide the following information:

- A review of the relevant published literature that supports development of transcriptomic PODs and their concordance with apical toxicological responses across a broad range of exposure routes, exposure durations, physicochemical properties, toxicokinetic half-lives, tissues, chemical classes, and technology platforms.
- Development of transcriptomic PODs for ETAP including study design and transcriptomic platform-specific refinement of the transcriptomic dose response methods based on:
    - concordance of the transcriptomic PODs with apical PODs from chronic, repeated dose toxicity studies;
    - inter-study variability of the transcriptomic PODs; and
    - family-wise error rates[13].
- Expected range of concordance between the transcriptomic and apical PODs in the context of the inter-study variability in both chronic, repeated-dose toxicity studies and short-term transcriptomic studies.

The second document details the methods for developing an ETAP and its implementation (EPA 2024b). That document is organized to provide the following information:

- Standard methods for developing an ETAP including the detailed methods for the initial database searches, systematic evidence mapping, short-term *in vivo* transcriptomic studies, human equivalent dose conversion, derivation of TRVs, and reporting and review.
- Comparison of TRVs with available EPA Integrated Risk Information System (IRIS) and Provisional Peer Reviewed Toxicity Value (PPRTV) reference doses (RfDs).
- ETAP reporting template.

[13] Family-wise error rate refers to the probability of at least one Type I error (defined as rejecting the null hypothesis, given that it is true).

3. REVIEW OF TRANSCRIPTOMIC DOSE RESPONSE LITERATURE

The literature summary provides expert-selected and reviewed studies that are relevant to application of transcriptomics in quantitative human health assessment in the specific context of the ETAP. The summary includes a brief history of transcriptomics at EPA, evaluation of the reproducibility of the technology, and the transcriptomic dose response literature as it pertains to the concordance of transcriptomic BMD(L)[14] with apical endpoints in standard toxicity studies. The literature search strategy is provided in Section 6.3 of the Appendix.

3.1. EPA HISTORY, POLICIES, AND USE OF TRANSCRIPTOMICS

The EPA has a longstanding commitment toward utilizing emerging and novel technologies to enhance testing paradigms and to improve the utility and predictability of risk assessment methods. Transcriptomic data are particularly attractive in this context because changes in gene transcript (mRNA) expression are frequently observed to precede or coincide with clinical or phenotypically observable changes (*i.e.,* apical effects), thereby providing a sensitive measurement of chemical-induced bioactivity (Hester et al. 2015; Lobenhofer et al. 2004; Thomas et al. 2011). Toward adoption of transcriptomics data, the Agency issued an interim policy on genomics in 2002 that advocated using transcriptomics data to enhance assessments and priority setting on a case-by-case basis within weight of evidence-based approaches (EPA 2004b). This Interim Policy was followed by a white paper issued by a task force of cross-Agency experts to investigate potential implications of transcriptomics on EPA programs and policies. While the Interim Policy limited consideration of transcriptomics data only in concert with traditional toxicology endpoints, the 2004 white paper on *Potential Implications of Genomics for Regulatory and Risk Assessment Application at the EPA* expanded the recommendations to include four areas of likely impact for genomics data within the Agency and for external submissions to the Agency. The recommended applications of transcriptomics data included prioritization of contaminants and contaminated sites and risk assessment (EPA 2004b). Key gaps that were identified as barriers to implementation at the time included a lack of adequate technical infrastructure and training of qualified personnel, as well as a need to develop a technical framework for genomic data analysis and standardization criteria for acceptance of transcriptomics data. In the same year (2004), the EPA's Office of Pesticides Programs utilized transcriptomics data in a mode of action, weight of evidence cancer risk assessment of

[14] The BMD(L) abbreviation is used in this document to define the combination of benchmark dose (BMD) and/or benchmark dose lower confidence bound (BMDL).

acetochlor (EPA 2004a). Supporting the mode of action assessment were time course-derived transcriptomic data of the rat olfactory mucosa at a single dose; early gene expression changes were consistent with oxidative damage to DNA followed by cell proliferation, and late gene expression changes were consistent with tumorigenic progression.

In 2007, the EPA released interim guidance for microarray data submissions, quality, and analysis – *Interim Guidance for Microarray-Based Assays; Data Submission, Quality, Analysis, Management and Training Considerations* (EPA 2007). The guidance was developed by a team of experts from multiple EPA programs and regional offices who were convened at the request of the EPA's Office of the Science Advisor (Hester et al. 2015). The interim guidance provided recommendations on performance approaches for quality assessment parameters, data analysis approaches, Agency data submissions, and data management practices. In addition, the guidance included a draft Genomics Data Evaluation Records template and recommendations for development of training modules and materials for risk assessors, cross-Agency collaboration, and case study applications.

Research into application of transcriptomic data to risk assessment continued to expand in subsequent years. In 2009, EPA released a case study for application of transcriptomic data to human health risk assessment of dibutyl phthalate (EPA 2009). The case study outlined a pragmatic and flexible approach to accommodate different health and risk assessment practices and focused primarily on informing mode of action as part of a weight of evidence. At the time, the authors noted limitations that resonated from prior guidance documents and case studies, specifically that there was a need for consistency in methods for interpreting and analyzing transcriptomics data.

For application to quantitative human health risk assessment, standardized methods and software applications have since been developed for analyzing transcriptomic changes as a function of dose through calculating BMD values at the gene, pathway, biological process, or molecular function level (Farmahin et al. 2017; NTP 2018; Phillips et al. 2019; Thomas et al. 2007; Yang et al. 2007). BMDExpress has become a widely adopted graphical user interface-based software package that facilitates analysis of transcriptomic dose response data (Phillips et al. 2019; Yang et al. 2007). BMDExpress employs validated continuous parametric models deployed in the EPA's BMDS software[15]. In the BMDExpress workflow, curve fits are computed for each gene transcript, followed by functional classification analysis that assigns transcriptomic features into pre-defined gene sets[16] [*e.g.*, Gene Ontology (GO) classes] and determines gene set level potency estimates for each of the populated gene sets [*i.e.,* a gene set BMD(L)]. The individual probe/gene-level modeling process is consistent with recommendations by the EPA for benchmark dose modeling in toxicological assessments (EPA 2012). The BMDExpress software release (version 2) is in use by EPA and partner

[15] EPA BMDS software can be found at: https://www.epa.gov/bmds
[16] A gene set is defined as pathway, biological process, molecular function, or other biologically based set of genes.

agencies, including Health Canada and the US National Institutes of Health (NIH), who developed the software along with the contractor Sciome (Phillips et al. 2019). Methods and assumptions underlying use of BMDExpress for analysis of transcriptomic dose response data were reviewed by an NIH-assembled panel of external experts with the resulting method published in 2018 (NTP 2018).

In parallel, standardized reporting templates for transcriptomics data have been developed in collaboration with international partners, addressing the remaining key gap identified in prior Agency guidances and case studies. In 2021, EPA scientists co-led an OECD team that created the first formal reporting framework of the processing and analysis of transcriptomic data for regulatory toxicology use [*i.e.*, the OECD Omics Reporting Framework[17] (Harrill et al. 2021b)]. Similarly, the NIH's Division of Translational Toxicology (DTT; formerly called the Division of the National Toxicology Program) has established standard methods and a reporting framework for their internal transcriptomic dose response studies and has used this reporting standard to disseminate data on several organophosphate flame retardant chemicals (NIEHS 2022b, c, d). Taken together, the past decade has yielded significant advances in the standardization of methods, analysis frameworks, and reporting standards, as well as training and capacity building, to address prior Agency-identified critical gaps and to support the use of transcriptomics in regulatory science at the EPA.

3.2. REPRODUCIBILITY OF TRANSCRIPTOMIC DATA

Confidence in the application of transcriptomics data is bolstered by the work of international consortia who have completed projects investigating inter- and intra-platform reproducibility of gene expression measurements by standard methods, including microarrays and ribonucleic acid sequencing (RNA-seq) technologies. Reliability of the methods used to measure changes in gene expression is one of the main principles identified as necessary for application of transcriptomic data in quantitative human health assessments (Johnson et al. 2022). The MicroArray Quality Consortium (MAQC; led by the US Food and Drug Administration) conducted cross-site, cross-platform studies to evaluate performance through the titration of two reference RNA samples (MAQC 2006). The Consortium demonstrated strong inter- and intra-platform reproducibility of gene expression measurements by microarrays and high correlation in quantitative expression between traditional reverse transcription-polymerase chain reaction (RT-PCR) assays and microarray results. The MAQC studies demonstrated that the combination of fold-change ranking and a non-stringent p-value cutoff led to increased consistency in differential gene expression results across laboratories. These and subsequent studies led to the MAQC launching the Sequencing Quality Control (SEQC) project to evaluate emerging next-generation sequencing (NGS) technologies. The first SEQC projects established best-practices for RNA-seq methods of measuring gene expression, compared RNA-seq

[17] The guidance document and template for the OECD Omics Reporting Framework can be found at: https://www.oecd.org/chemicalsafety/testing/omics.htm

performance to microarrays (SEQC/MAQC-III 2014), characterized inter-platform reproducibility of RNA-seq protocols and technologies (Li et al. 2014), and evaluated the bioinformatic tools increasingly required to analyze large and complex RNA-seq datasets (Li et al. 2014). Similar to the microarray reproducibility studies, the SEQC cross-site, cross-platform study demonstrated high consistency in RNA-seq results with intra- and inter-platform Spearman rank concordance values at $r > 0.86$ and $r > 0.83$, respectively. (Li et al. 2014).

In 2021, the MAQC published results of its largest project to date, known as SEQC2 (Sequencing Quality Control Phase 2) (Mercer et al. 2021). SEQC2, funded by the 21st Century Cures Act, benefited from international participation of more than 300 contributing scientists from 150 industry, academic, and government organizations. It was divided into six areas of clinical application which used a variety of NGS technologies (DNA-seq, RNA-seq, Exome-seq, methylation-seq, etc.) for germline variant detection, cancer genomics, cell-free tumor DNA/biomarker discovery, and other precision medicine initiatives. Among the goals of SEQC2 toward analytical validation of transcriptomic data streams were to: 1) develop tangible and reproducible reference standards for NGS platforms; 2) benchmark the impact of experimental and bioinformatic variables on the generation and analysis of NGS data; and 3) evaluate inter- and intra-lab reproducibility of transcriptomic data derived from NGS technologies across different laboratories (Foox et al. 2021). The SEQC2 project significantly advanced understanding of the translational scientific infrastructure afforded by NGS technologies. The SEQC2 project yielded the scientific foundation that led to U.S. Food and Drug Administration (FDA) approvals of a number of NGS-based solid tumor tests as companion diagnostics (Li et al. 2021) utilized in patient therapy and precision oncology[18]. Most recently, FDA has approved NGS-based companion diagnostic (CDx) tests for the monitoring of advanced solid tumors using liquid biopsies, *e.g.,* Guardant360® CDx[19], and FoundationOne® Liquid CDx. Use of RNA-seq technologies in patient care settings demonstrates the degree of confidence by federal regulatory agencies, such as FDA, in utilizing reproducible and validated NGS workflows to advance human health protection.

3.3. DOSE CONCORDANCE OF TRANSCRIPTIONAL AND APICAL RESPONSES

Traditionally, subchronic or chronic duration toxicity studies, such as the two-year rodent bioassay, are conducted to inform the qualitative landscape of potential adverse health effects of environmental chemicals and to identify quantitative PODs considered in the derivation of reference values in human health assessments. However, given the general lack of traditional toxicity testing

[18] The list of FDA approved nucleic acid tests can be found at: https://www.fda.gov/medical-devices/in-vitro-diagnostics/nucleic-acid-based-tests#top

[19] The news release can be found at: https://www.fda.gov/news-events/press-announcements/fda-approves-first-liquid-biopsy-next-generation-sequencing-companion-diagnostic-test

data and human health assessments for environmental chemicals, shorter duration animal studies combined with large scale molecular measurement technologies provide a significant opportunity to fill these gaps. Time course studies indicate that upon sustained exposure to stimuli, such as environmental chemicals, changes in cellular gene expression can occur within minutes to days (Tullai et al. 2007). Frequently, progression of gene expression changes following chemical exposure involves induction of genes associated with xenobiotic response, such as nuclear transcription factors that are critical mediators of a wide variety of key biological functions (*e.g.,* stress responses, immune effects, and metabolic changes) that result in downstream protein production and physiologic responses, including adverse health effects [reviewed in (Fowler et al. 2011)].

Quantifying transcriptional changes provides an indication of chemical potency for producing biological perturbation(s). Indeed, multiple studies have evaluated the relationship between transcriptomic BMD(L) values and traditional tissue-, organ-, and organism-level BMD(L) values for adverse apical endpoints across a variety of exposure durations, treatments, routes of exposure, biological targets, and species [*e.g.*, (Bercu et al. 2010; Chepelev et al. 2015; Clewell et al. 2011; Clewell et al. 2014; Dunnick et al. 2017; Gwinn et al. 2020; Jackson et al. 2014; Johnson et al. 2020; Labib et al. 2016; Lake et al. 2016; Recio et al. 2017; Shockley et al. 2020; Thomas et al. 2007; Thomas et al. 2012; Zhou et al. 2017)]. Collectively, these studies and others reviewed in this document demonstrated that transcriptomic BMD(L) values, when grouped by pathway, biological process, molecular function, or other biologically based sets of genes, showed good concordance with BMD(L) values for phenotypic apical effects from traditional animal-based toxicity studies. The ability to derive a BMD(L) value that indicates a coordinated gene expression response and its association with an adverse toxicological response are among the main principles necessary to apply transcriptomic data to quantitative human health assessments (Johnson et al. 2022).

3.4. IMPACT OF EXPOSURE DURATION ON TRANSCRIPTIONAL AND APICAL DOSE CONCORDANCE

One consideration for quantifying a transcriptomic POD from short-term *in vivo* studies is to determine the duration of exposure needed to inform adverse health effects that may develop in longer-duration (*e.g.*, chronic) studies. Studies conducted across several organizations (*i.e.,* EPA, Hamner Institutes, Health Canada, NTP, Corteva, and others) have experimented with different timepoints – ranging from 1 to 90 days – with the goal of comparing dose-concordance between gene set-based transcriptomic PODs and apical effect-based PODs from traditional toxicity test designs/exposure durations. One of the aims of these investigations has been to identify an "optimal" repeated dosing duration for transcriptomic assessment.

Multiple studies have compared apical effect and transcriptomic BMD values in subchronic studies (Bianchi et al. 2021; Clewell et al. 2011; Clewell et al. 2014; Dong et al. 2016; Thomas et al. 2011). In 2011, Thomas and colleagues exposed female mice for 90 days via gavage or inhalation to

five chemicals[20] that had been previously evaluated in conventional subchronic and/or chronic duration studies (*e.g.*, by the NTP). Gene set-based transcriptomic BMD values from the target tissues (*i.e.*, liver or lung) were compared to 90-day and two-year gross tissue and histopathology data. Both tissues showed concordance between the transcriptomic BMD(L) and the lowest apical BMD(L) values associated with histological or tissue weight change at those timepoints. All transcriptomic BMD values were within 8-fold (10-fold for the transcriptomic BMDL) of the lowest apical BMD value for both non-cancer and cancer effects. While overall the transcriptomic BMD values tended to be lower than apical effect values, there were instances where the opposite was true. In those cases, the transcriptomic BMD was no more than 3-fold higher than the apical effect BMD (Thomas et al. 2011). In another subchronic study, Dong and colleagues identified a transcriptomic BMD for furan in male rat liver tissue that was slightly higher (2-fold) than the time-matched BMD for the lowest histopathological endpoints of Kupffer cell pigmentation and hepatocyte apoptosis (Dong et al. 2016). When compared to two-year adenoma and carcinoma incidence data, the transcriptomic BMD was 30-fold more sensitive than the tumor-related BMD. Insufficient gene response in females prevented a similar comparison, which also coincided with prior observations of females being generally less sensitive to furan exposure (Dong et al. 2016). More recently, Bianchi and colleagues compared gene set-based transcriptomic BMD(L) values from male rats exposed to four agrochemicals[21] for 90-days to the pathological effects at that time-point. Gene expression data were obtained from liver, kidney, or, in one case, both tissues, and transcriptomic PODs were compared to the apical, 90-day rat and two-year rat and mouse data regardless of target tissue or sex (Bianchi et al. 2021). When comparing species-matched results, the male transcriptomic POD was 1.2- to 14.1-fold lower than the apical effect POD for either sex in rats at 90 days. The subchronic transcriptomic POD values performed slightly less well when correlated with the two-year rodent data; yet these values were 1.3- to 22.9-fold lower (for all but one chemical, sulfoxaflor) than the apical POD. On average, the rat transcriptomic PODs tended to be within 10-fold of the lowest two-year, apical POD (Bianchi et al. 2021). Overall, these studies indicate that transcriptomic PODs provide reasonable estimates of apical effect PODs following subchronic exposures.

A larger proportion of studies have focused BMD modeling of transcriptomics from shorter-duration exposures (3- to 30-days) to inform carcinogenic mode of action and identify PODs (Bhat et al. 2013; Chepelev et al. 2017, 2018; Jackson et al. 2014; Labib et al. 2016; Labib et al. 2017; Moffat et al. 2015). Several studies had time-matched clinical and histopathological data for comparison with gene set-based transcriptional responses with many focused on the 28- or 30-day timepoint. For instance, a case study by Bhat and colleagues reported four tumorigenic conazoles (cyproconazole, epoxiconazole, propiconazole, and triadimefon) tested in male mice were

[20] The five chemicals were: 1,4-dichlorobenzene; 1,2,3-trichloropropane; propylene glycol mono-*t*-butyl ether; naphthalene; and methylene chloride.
[21] The four chemicals were triclopyr acid; sulfoxaflor; pronamide; and fenpicoxamid.

differentiated from a non-tumorigenic conazole (myclobutanil) after 30 days of exposure with transcriptomic data, but traditional pathology data (liver weight, tumor burden) could not distinguish tumorigenic from non-tumorigenic conazoles at the 30-day time point (Bhat et al. 2013). The 30-day gene set-based transcriptomic POD values were the same as or within 2-fold of the time matched liver weight changes. For the tumorigenic conazoles, the gene set transcriptomic POD values were within 5-fold of the apical PODs based on hepatocellular carcinoma or adenoma for each chemical, which ranged from 2.1 mg/kg-day for cyproconazole to 270 mg/kg-day for triadimefon (Bhat et al. 2013). In another study that investigated furan, Jackson and colleagues demonstrated that transcriptomic PODs after 3 weeks of exposure in female mice were within 12-fold of the 90-day apical POD for hepatocyte cell death and equivalent to the two-year apical POD for hepatocellular adenoma (Jackson et al. 2014).

For the few studies with time-matched clinical and histopathological data at multiple early timepoints, there are indications that shorter studies are just as sensitive for determining apical PODs as longer duration studies. For instance, Moffat and colleagues explored the added value of combining transcriptomic data with traditional toxicity data in conducting a risk assessment of benzo(a)pyrene. Gene set level response data from 3- and 28-day mouse were related to lung, forestomach, and liver cancer development by either focusing on key events in carcinogenesis or using no *a priori* toxicity knowledge and using enriched gene sets (Moffat et al. 2015), similar to the approach employed by Thomas and colleagues (2011) except the 10th percentile of all genes assigned to an Ingenuity Pathway Analysis gene set BMD was used instead of the median. Both the Moffat and Thomas approaches produced liver-based transcriptomic PODs after 3 days of exposure that were lower than the 28-day transcriptomic PODs by 6.7- and 26.5-fold for the key event and no *a priori* knowledge approaches, respectively, with the caveat that the 28-day transcriptomic data were obtained 3 days after the final exposure. The 3-day transcriptomic PODs were also lower compared to the chronic apical POD for liver tumors at 1.2 mg/kg-day. However, the gene set transcriptomic POD in the no *a priori* knowledge approach was slightly lower at 0.2 mg/kg-day than the transcriptomic POD using mode of action information at 1 mg/kg-day suggesting either approach could be useful for assessing the risk of chemicals that lack traditional toxicity testing data (Moffat et al. 2015). Another two studies in mice and rats investigating acrylamide-induced tumor development collected transcriptomic dose response data at 5, 15 and 31 days for comparison to two-year tumor data in the Harderian gland (the target tissue with the lowest apical POD), lung, and forestomach in mouse or the liver and thyroid (the target tissue with the lowest apical POD) in rat (Chepelev et al. 2017, 2018). Chepelev and colleagues (2017 and 2018) identified that 15 days was sufficient for identifying a transcriptomic POD consistent with thyroid and lung tumors for rat and mouse, respectively. Both were slightly higher than the apical PODs at 2.1- and 1.4-fold; however, there were insufficient gene expression changes after 5 days of exposure to identify a transcriptomic POD in either study, and 31 days of exposure were required to identify a transcriptomic POD from the Harderian gland which was within 2.3-fold of the apical POD for tumor development in the mouse

(Chepelev et al. 2018). A related study focusing on acrylamide-related testicular toxicity also evaluated gene response data in the testes at 5, 15 and 31 days and found positive dose-related trends on days 5 and 31; however, only day-31 genes underwent transcriptomic BMD modeling. From this, a transcriptomic POD was identified for testes that was slightly lower than the No Observable Adverse Effect Level (NOAEL) for male reproductive toxicity, but was 6-fold higher than the BMDL values for other apical effects such as peripheral neuropathy (Recio et al. 2017). Gwinn and colleagues evaluated acrylamide in a 5-day transcriptomic study and found that the lowest median BMD value when summarized by gene set was equivalent to the lowest apical BMD for peripheral nerve degeneration in male rats (0.68 and 0.61 mg/kg-day, respectively)(Gwinn et al. 2020). Gwinn and colleagues also identified Harderian gland adenoma and adenocarcinoma as the apical effect with the lowest POD, which was less than 2-fold lower than the gene set level BMD (Gwinn et al. 2020). The difference between the Gwinn and Recio studies may be due to the tissue used for measuring gene response. Gwinn and colleagues used kidney and liver, which may be more indicative of systemic effects and could explain the lower transcriptomic BMD compared to the testes gene response from the Recio study.

A recent study by Johnson and colleagues examining the consistency in transcriptomic PODs across multiple timepoints provides support for using 5- to 30-day exposure durations to inform adverse effect PODs observed at timepoints beyond 30 days (Johnson et al. 2020). In this work, TG-GATES[22] data were analyzed to compare transcriptomic and apical PODs across 79 chemicals for both a single dose and repeated dose study designs. Transcriptomic POD values from the liver of exposed male rats were used as the sentinel tissue to compare with the lowest apical PODs at the 29-day timepoint. The apical endpoints examined included histopathology in liver and kidney, body and organ weights, and clinical observations. The transcriptomic PODs were generally consistent across later timepoints (>4 days) relative to the 29-day apical POD values regardless of dosing scheme (r > 0.81 for the 4-, 8-, and 15-day timepoints). For the 51 chemicals with repeated dose data at the later timepoints, >90% had transcriptomic PODs within 10-fold of the lowest 29-day apical POD. Across dosing regimens and timepoints, the median absolute fold change in transcriptomic POD relative to the 29-day apical POD was approximately two (Johnson et al. 2020).

In one of the more thorough time course studies, Thomas and colleagues assessed the transcriptional response from six[23] chemicals at multiple time points (5 days and 2, 4, and 13 weeks)

[22] TG-GATEs (Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System) is a public transcriptomic genomic database resulting from two joint government-private sector projects (H14-Toxico-001 and H19-Toxico-001) organized by the Japanese National Institute of Biomedical Innovation, National Institute of Health Sciences, and multiple pharmaceutical companies. The project generated gene expression and toxicity data in rats and primary cultured hepatocytes of rats and humans following exposure to 170 compounds in dose response, of which 79 compounds that met filtering criteria were utilized by Johnson et al. (2020).

[23] The six chemicals were 1,2,4-tribromobenzene; bromobenzene; 2,3,4,6-tetrachlorophenol; 4,4'-methylenebis (N,N-dimethyl) benzenamine, N-nitrosodiphenylamine, hydrazobenzene.

in male or female rats across a range of target tissues (liver, thyroid, and bladder) and two oral exposure administration methods (gavage and feed) (Thomas et al. 2013b). The lowest transcriptomic BMD values when summarized by gene set remained stable across time. The transcriptomic BMD values correlated robustly with non-cancer phenotypic responses across all exposure durations from 5 days, 2, 4, and 13 weeks ($r > 0.9$) (Thomas et al. 2013b). For cancer-related effects, the transcriptional and tumor BMD values were also highly concordant. A reanalysis of the Thomas et al. 2011 and 2013 datasets compared 11 different approaches for calculating transcriptomic BMDs and identified multiple gene set and gene-based approaches that were in agreement with the traditional non-cancer and cancer apical BMD values (Farmahin et al. 2017). A slightly older study examining the time-course effects of paraformaldehyde across 1-, 4-, and 13-week exposure periods also demonstrated that the transcriptomic POD values from short-term exposures were only slightly higher than the subchronic transcriptomic POD, and both values were within 1.4-fold of the NOAEL for nasal tumor development (Andersen et al. 2010).

In a time course study examining the transcriptional and apical dose concordance in sensitive and non-sensitive species, Thomas and colleagues measured transcriptomic changes in female mouse and female rat lungs following 5 and 15 days of exposure to 2-chloro-1,3-butadiene (Thomas et al. 2013a). The transcriptional and apical concordance in BMDs and PODs was strong at both time points in the sensitive species (*i.e.*, female mice); however, in the less sensitive species (*i.e.*, female rat), the 5-day time point showed better dose concordance than the 15-day time point. At the 15-day time point in the female rat, the transcriptomic BMD and POD were lower than the apical BMD for lung tumors.

A significant number of studies have examined transcriptomic and apical dose concordance following exposures of 7 days or less. Lake and colleagues (2016) investigated the ability of transcriptomics to distinguish the differences in tumorigenic potency for three reference phthalates (diethylhexyl phthalate, di-n-octyl-phthalate, and butyl benzyl phthalate) with differing levels of peroxisome proliferator activated receptor α (PPARα) activation following 7 days of exposure. Transcriptomic POD estimates correctly stratified the phthalates according to tumorigenic potencies (Lake et al. 2016). Zhou and colleagues (2017) exposed mice to either trichloroethylene or tetrachloroethylene for 1-day and performed whole transcriptome RNA-Seq on livers and kidneys. In order to compare across species, they employed physiologically-based pharmacokinetic (PBPK) modeling to calculate human equivalent doses for mouse tissue-specific transcriptomic PODs as well as apical PODs from two-year rodent bioassays (Zhou et al. 2017). When considering established mode of action for these two chemicals, the median, human equivalent, transcriptomic PODs for liver and kidney were 10.4-fold (± 3.5 MAD) higher than the lowest, two-year, apical POD. Conversely, when the lowest median transcriptomic POD summarized by gene set was used for comparison, the values tended to be within 2.6-fold (median ± 1.2 MAD) of the lowest, two-year apical PODs (Zhou et al. 2017). Bercu and colleagues demonstrated that 2-day fenofibrate and up to 7-day methapyrilene exposures in female rats resulted in transcriptomic PODs that were either 12-fold lower or 8-fold

higher than the apical PODs for hepatocellular carcinoma for each pharmaceutical, respectively (Bercu et al. 2010).

Numerous studies performed by the NTP and the DTT have included short-term transcriptomic studies for comparison to studies of longer duration (Catlin et al. 2018; Dunnick et al. 2017; Gwinn et al. 2020; NIEHS 2022a; Shockley et al. 2020). Initial studies from DNTP using a 5-day transcriptomic assay examined N,N-dimethyl-p-toluidine and toluidine in male rats. Dunnick and colleagues developed transcriptomic PODs by taking the mean BMDL from gene sets that had at least 5 genes altered (Dunnick et al. 2017). While this study itself did not directly compare transcriptomic PODs to apical PODs, for N,N-dimethyl-p-toluidine, the gene set with the lowest transcriptomic POD (2.0 mg/kg-day) was three times higher than the lowest apical POD (0.64 mg/kg-day) for respiratory metaplasia in male rats from the NTP bioassay (NTP 2022). Shockley and colleagues examined a series of brominated or phosphorylated flame retardants in 5-day transcriptomic assays. Two of these chemicals, decabromodiphenyl ether and 2,2',4,4'-tetrabromodiphenyl ether, have IRIS values based on mouse neurobehavioral toxicity in which the apical POD used to derive the reference dose (RfD) is within a factor of 7 or less of the transcriptomic POD (Shockley et al. 2020). In a series of NIEHS reports, the DTT studied four phosphorylated flame retardants in 5-day transcriptomic studies. Of the four phosphorylated flame retardants, only tricresyl phosphate has a publicly available guideline study. The transcriptomic POD for tricresyl phosphate (10.1 mg/kg-day) (NIEHS 2022a) was approximately 2-fold lower than a guideline developmental toxicity study in rats[24] (20 mg/kg-day) found in the European Chemicals Agency's (ECHA) database.

In an expansive DNTP study, targeted gene expression measurements were made in the liver and kidneys of male rats following exposure to 18 chemicals or mixtures at 8 dose levels with a study duration of 5 days (Gwinn et al. 2020). The study used the S1500+ TempO-Seq transcriptomic platform (Mav et al. 2018). For the chemicals that had corresponding repeated dose toxicity studies in the same sex and species (*i.e.*, male rats), the gene set with the lowest median transcriptomic BMD value from the liver or kidneys was within 10-fold of the lowest non-cancer apical BMD for >90% of tested substances, including apical BMD values from other organs. For cancer apical BMDs, transcriptomic BMD values from the liver or kidneys were within 10-fold of the lowest cancer apical BMD for >80% of tested chemicals or mixtures (15/18 substances). Notably, the concordance in transcriptomic and apical BMD values was not sensitive to differences in the toxicokinetic half-lives. The chemicals in the DNTP study had half-lives in male rats of hours (acrylamide and bromodichloroacetic acid) to weeks (perfluorooctanoic acid and the pentabromodiphenyl ether mixture DE-71). In the Gwinn et al. study, transcriptional changes and transcriptomic values in liver and kidney were identified for two botanicals, ginseng and milk thistle extract, as well as tetrabromobisphenol A, which had shown non-cancer or cancer-;related histopathological lesions in

[24] ECHA REACH Dossier at: https://echa.europa.eu/registration-dossier/-/registered-dossier/16010

male rats in previously conducted two-year chronic bioassays (Gwinn et al. 2020). However, in the original NTP reports, tetrabromobisphenol A was reported to have a reduction in body weight at the middle and high dose levels (NTP 2014) and milk thistle extract showed a reduction in biliary hyperplasia and mixed cell infiltration at the high dose (NTP 2011). While the histopathological changes following exposure to milk thistle extract may not be considered adverse in a traditional assessment context, apical changes were present. When the chemicals or mixtures from the Gwinn et al. study were evaluated across species and sexes, the lowest male rat transcriptional BMD was within an order of magnitude of the lowest apical BMD in male and female rats and mice from the NTP chronic toxicity and carcinogenicity studies 82% of the time.

To illustrate the concordance of transcriptional BMD(L) from various exposure durations with apical BMD(L) from chronic toxicity and carcinogenicity studies, a scatter plot was created (Fig. 3-1). Studies included in the scatter plot and concordance analysis focused on those reporting both gene set-based BMD(L) values from 1- to 90-day exposures and chronic apical BMD(L). When all the studies were combined, the Pearson's correlation coefficient for the transcriptomic BMD versus chronic, apical BMD was 0.820 with a $\log_{10}$ RMSD of 0.593 ($\log_{10}$ mg/kg-day) and a median absolute ratio[25] of 2.4 ± 1.0 (MAD). For exposure durations of 5 days, the Pearson's correlation coefficient for the transcriptomic BMD versus chronic, apical BMD was 0.760 with a $\log_{10}$ RMSD of 0.621 ($\log_{10}$ mg/kg-day) and median absolute ratio of 2.7 ± 1.3 (MAD). However, the 5-day exposure duration data contained a larger mix of chemicals where transcriptomic changes were not measured in the same tissue as the chronic, apical endpoint. For comparison, the RMSD values for the combined dataset are similar to the range of inter-study standard deviation estimates for the Lowest Observable Adverse Effect Levels (LOAELs) for systemic toxicity in repeated dose studies, approximated as residual RMSE in log10-mg/kg-day units [0.45-0.56; (Pham et al. 2020)]. The results suggest that the error associated with the concordance between the transcriptomic BMD values versus non-cancer and cancer apical BMD values is approximately equivalent to the inter-study variability in the repeated dose toxicity study itself. Data for the comparison in Figure 3-1 were obtained from studies published between 2010 and 2022 (Andersen et al. 2010; Bercu et al. 2010; Bhat et al. 2013; Bianchi et al. 2021; Cannizzo et al. 2022; Chepelev et al. 2017, 2018; Dong et al. 2016; Gwinn et al. 2020; Hester et al. 2016; Jackson et al. 2014; Labib et al. 2017; LaRocca et al. 2020; Moffat et al. 2015; Recio et al. 2017; Thomas et al. 2007; Thomas et al. 2011; Thomas et al. 2013a; Thomas et al. 2013b; Zhou et al. 2009).

[25] The absolute ratio between a and b is defined as maximum{a/b, b/a}.

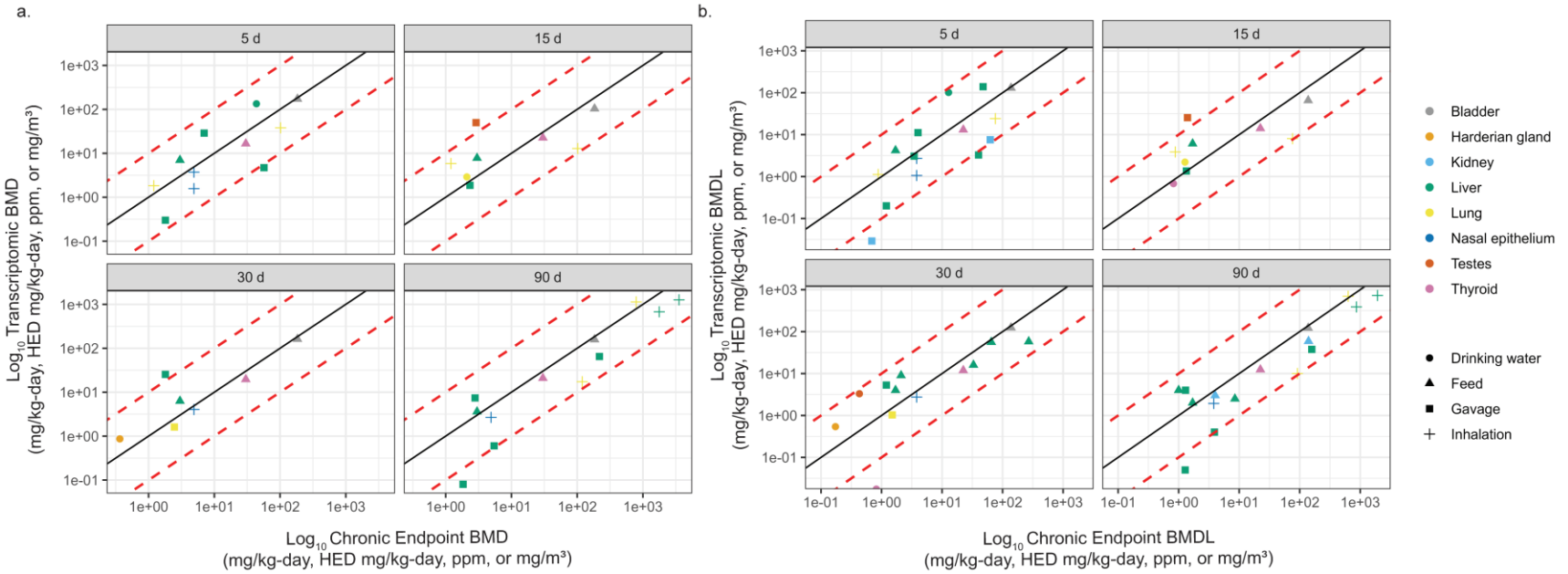


Figure 3-1. Concordance between mouse or rat transcriptomic and chronic apical BMDs or BMDLs including carcinogenicity studies. (a.) Transcriptomic BMD values vs. two-year apical BMD values. (b.) Transcriptomic BMDL values vs. two-year apical BMDL values. Transcriptomic results from 1, 2, 3, 5, 6, or 7-day test durations are reported under the 5-day timepoint. Results from 15 and 21-day test durations are reported under the 15-day timepoint. Results from the 28, 29, 30, and 31-day exposure durations are reported under the 30-day timepoint, and results from the 90 and 91-day exposure durations are reported under the 90-day timepoint. Only studies with a transcriptomic BMD and/or BMDL and two-year apical BMD and/or BMDL were included in the graphs. For the transcriptomic BMD and/or BMDL, the median of the most sensitive gene set or datapoint most similar to the ETAP approach was used. If multiple tissues were investigated in the study, then the lowest transcriptomic BMD(L) across all tissues was selected. For the apical BMD(L), the lowest two-year BMD(L) reported for a chemical and species was used, which sometimes did not match the tissue used for obtaining the transcriptomic BMD(L). If the same chemical was independently tested in different studies, then the transcriptomic BMD(L) and associated apical BMD(L) for each study was included. Only dose matched metrics were used. For instance, if the apical BMD(L) was in mg/kg-day, then the transcriptomic BMD(L) reported in mg/kg-day was used, but if the apical BMD(L) was in ppm, then the transcriptomic BMD(L) reported in ppm was used. The header for each facet of the figure indicates the timepoint in days (d) for the transcriptomic BMD(L). HED is human equivalent dose. The solid line indicates a perfect concordance at 1. The dashed lines represent 10-fold difference between the BMD(L) values.

3.5. IMPACTS OF CHEMICAL MODE-0F-ACTION ON TRANSCRIPTIONAL AND APICAL DOSE CONCORDANCE

The studies examining transcriptomic PODs in shorter-term exposure durations typically focused on how well they predict apical PODs from chronic, subchronic and shorter-term toxicity studies. The chemicals evaluated cover a broad set of bioactivities. For example, Johnson and colleagues examined 79 chemicals, most of which were pharmaceuticals. Included in this list were neuroactive chemicals (*e.g.*, diazepam, phenobarbital, haloperidol), anticancer agents (*e.g.*, doxorubicin, etoposide, cyclophosphamide), antibiotics (*e.g.*, isoniazide, tetracycline), and several other classes of drugs (Johnson et al. 2020). These chemicals demonstrate a broad range of molecular initiating events that cover receptor-mediated responses, enzyme inhibition and DNA alkylation. In contrast, the chemicals examined by Thomas and colleagues (2013) were mainly industrial chemicals with unknown modes of action. The studies by the DNTP (Dunnick et al. 2017; Gwinn et al. 2020; Shockley et al. 2020) examined genotoxic chemicals (*e.g.*, acrylamide, furan, N,N-dimethyl-p-toluidine), receptor-mediated and endocrine-active toxicants (*e.g.*, perfluorooctanoic acid, ethinyl estradiol; di(2-ethylhexyl)phthalate) as well as chemicals with uncertain modes of action (*e.g.*, 3.3',4,4'-tetrachlcoroazobenzene, hexachlorobenzene, bromodichloroacetic acid, decabromodiphenyl ether). While there has not been a systematic approach to evaluating transcriptional and apical dose concordance across specific molecular initiating events or modes of action, the diversity of chemicals evaluated in these studies has covered sufficient biological space to suggest that gene set-based transcriptomic PODs from short-term *in vivo* studies would work well for most general and target organ toxicities of interest to risk assessors.

Currently, studies evaluating the concordance of transcriptomic PODs with apical endpoints have primarily focused on subchronic and chronic toxicity studies. Additional studies may be required to evaluate how well transcriptomic PODs from adult animals predict apical PODs for other health domains, such as developmental toxicity, reproductive toxicity, immunotoxicity, and neurotoxicity. Notably, a previous retrospective analysis demonstrated that apical data from subchronic toxicity studies can predict effects on fertility (Dent 2007), while a separate retrospective analysis showed less than a two-fold difference in NOAEL values between rat subchronic studies and two-generation reproductive and developmental studies (Janer et al. 2007). Given the concordance between transcriptomic PODs from short-term studies and apical PODs from subchronic and chronic toxicity studies, these two retrospective analyses suggest that transcriptomic PODs may also be protective of apical PODs for other toxic effects.

Despite multiple efforts to relate gene set alterations to apical responses, there are few toxicities for which there is a consensus for causal relationships between gene set changes leading to these apical responses. It should be noted that, in studies examining the time course of transcriptional changes, different gene sets were altered at different time points, yet the transcriptomic BMD(L) values remained generally consistent over time (Johnson et al. 2020; Thomas et al. 2013b).

3.6. SCOPE OF CHEMICAL FUNCTIONAL USE, PROPERTIES, AND TOXICOKINETIC PARAMETERS FOR CHEMICALS WITH TRANSCRIPTIONAL AND APICAL DOSE CONCORDANCE

The chemicals evaluated in short-term transcriptomic studies have included pharmaceuticals (Bercu et al. 2010; Gwinn et al. 2020; Johnson et al. 2020), pesticides (Bhat et al. 2013), flame retardants (Gwinn et al. 2020; Shockley et al. 2020), disinfection by-products (Gwinn et al. 2020; Shockley et al. 2020), persistent organic pollutants (Dunnick et al. 2017; Gwinn et al. 2020), high production volume chemicals (Gwinn et al. 2020; Thomas et al. 2013b), solvents (Catlin et al. 2018; Gwinn et al. 2020) and botanicals (Catlin et al. 2018; Gwinn et al. 2020). These substances not only have diverse commercial uses, but also encompass a large range of physicochemical properties. For comparison, a subset of physicochemical properties was compared between the 140 chemicals in the transcriptomic studies in the literature review and the chemicals on the TSCA active inventory; however, not all chemicals on the TSCA active inventory have defined structures, nor are physicochemical properties able to be predicted for all chemicals. Nonetheless, while the full range for each property was typically larger in the TSCA active inventory, the 10th and 90th percentiles were similar for both sets of chemicals suggesting that most of the distribution in properties were similar (Table 3-1). Apart from the physicochemical properties, the toxicokinetic parameters also vary significantly in this group of chemicals. For example, many of the pharmaceuticals evaluated have elimination half-lives of hours (caffeine, ethanol) (Johnson et al. 2020), while some of the industrial chemicals studied have elimination half-lives of months in rodents (Gwinn et al. 2020; Shockley et al. 2020). While many of the pharmaceuticals studied are extensively metabolized, some of the industrial chemicals such as perfluorooctanoic acid and decabromodiphenyl ether are poorly metabolized, if at all. The range of physicochemical and toxicokinetic properties of the chemicals studied suggests that short-term transcriptomic studies can provide reasonable estimates of apical PODs for a broad range of chemicals and substances.

Table 3-1. Predicted physicochemical properties of the chemicals evaluated in the *in vivo* transcriptomic studies in comparison with chemicals on the TSCA active inventory[a].

Physicochemical Property	Median	Minimum	Maximum	10th Percentile	90th Percentile
Chemicals Evaluated in *In Vivo* Transcriptomic Studies					
Molecular Mass (g/mol)	272.04	46.07	1202.64	132.12	426.12
LogK$_{ow}$	2.85	-1.78	8.10	-0.29	5.76
Vapor Pressure (mmHg)	3.18e-07	1.53e-11	5.95e02	5.36e-10	1.34e00

Table 3-1. Predicted physicochemical properties of the chemicals evaluated in the *in vivo* transcriptomic studies in comparison with chemicals on the TSCA active inventory[a].					
Henry's Law (atm-m3/mol)	4.66e-08	6.84e-12	1.77e-02	1.39e-10	3.23e-04
Water Solubility (mol/L)	4.00e-04	1.05e-10	3.77e01	8.54e-07	3.27e-01
Chemicals on the TSCA Active Inventory					
Molecular Mass (g/mol)	221.03	2.02	4506.92	117.15	558.19
$LogK_{ow}$	2.38	-5.08	10.28	-0.9	5.64
Vapor Pressure (mmHg)	2.78e-04	3.06e-14	2.84e09	9.78e-10	5.26e00
Henry's Law (atm-m3/mol)	2.57e-07	2.53e-12	3.29e01	2.19e-10	5.43e-04
Water Solubility (mol/L)	2.79e-03	6.84e-14	3.77e01	1.26e-06	1.19e00
[a]Predicted physicochemical properties were obtained from the EPA CompTox Chemicals Dashboard using the OPERA model. Physicochemical properties were not able to be predicted for all chemicals.					

3.7. IMPACTS OF ROUTE OF EXPOSURE AND TISSUE SELECTION ON TRANSCRIPTIONAL AND APICAL DOSE CONCORDANCE

Common routes of chemical exposure in traditional toxicity studies include inhalation, ingestion, and absorption through the skin and eyes. Studies that have derived transcriptomic PODs have generally used oral exposures including feed, gavage, and drinking water. For the inhalation route, Thomas and colleagues used inhalation exposures for 3 chemicals (propylene glycol mono-t-butyl ether, methylene chloride, and naphthalene) (Thomas et al. 2011) and Andersen and colleagues performed time course inhalation studies using formaldehyde (Andersen et al. 2008). In addition, Thomas and colleagues evaluated cross-species concordance in transcriptomic PODs following inhalation exposure to 2-chloro-1,3-butadiene for 5 and 15 days (Thomas et al. 2013a). No studies were identified that evaluated transcriptomic PODs for chemicals in which dermal or ocular exposures were employed. For inhalation and oral exposures, the available studies suggest that the route of exposure does not appear to significantly alter the relationship between the transcriptomic and apical POD (Fig. 3-1). For oral studies across all exposure durations, the Pearson's correlation coefficient for the transcriptomic BMD versus chronic, apical BMD was 0.762 with a $\log_{10}$ RMSD of 0.618 and median absolute ratio of 2.2 ± 0.9 (MAD). For inhalation studies, the Pearson's correlation coefficient for the transcriptomic BMD versus chronic, apical BMD was 0.933 with a $\log_{10}$ RMSD of 0.495 and median absolute ratio of 2.7 ± 1.2 (MAD).

In guideline chronic toxicity and carcinogenicity studies, approximately 40 different tissues from rats are typically selected for histopathological analysis. Transcriptional analysis of all 40 tissues in male and female rats across an expanded dose range would be very costly. To control costs, the majority of the published studies evaluating the transcriptomic and apical dose concordance focused only on a few selected target tissues as defined by endpoints in subchronic or chronic guideline studies. A subset of studies across 19 chemicals have evaluated use of liver and/or kidney as sentinel tissues for determining transcriptomic PODs (Bianchi et al. 2021; Gwinn et al. 2020; Thomas et al. 2011; Thomas et al. 2012; Zhou et al. 2017). The liver and kidneys were chosen as sentinel tissues for several reasons. First, the liver is the principal site of xenobiotic metabolism immediately after chemical absorption from the gastrointestinal tract, and it has the largest supply of biotransformation enzymes (*e.g.,* cytochromes P450) of all organs in the body. Therefore, the liver has a key role in xenobiotic detoxification. For some xenobiotics, hepatic metabolism may result in activation of the chemical. For example, the carcinogen benzo(a)pyrene must be oxidized to its ultimate carcinogenic form, a diol epoxide. Thus, the liver is a key sentinel tissue for transcriptomic analysis due to its role in xenobiotic metabolism. Second, the kidney is frequently selected as a sentinel tissue because it receives about 20% of the cardiac output. The kidney functions to remove xenobiotics and their metabolites from the blood and excrete them from the body through urine. Finally, the kidney can also transport chemicals into the tubular lumen of the kidney where absorption from the lumen can also occur. Across the 19 chemicals that have been studied using liver and kidney as sentinel tissues, the combined data from these studies show slightly lower concordance between the transcriptomic and apical PODs (Fig. 3-2). The Pearson's correlation coefficient for the transcriptomic BMD versus chronic, apical BMD for the sentinel tissues was 0.797 with a $\log_{10}$ RMSD of 0.678 and median absolute ratio of 3.6 ± 2.0 (MAD). However, from a cost and efficiency standpoint, the results demonstrate that the use of a smaller set of sentinel tissues can provide predictions of the apical PODs within approximately ± 10-fold.

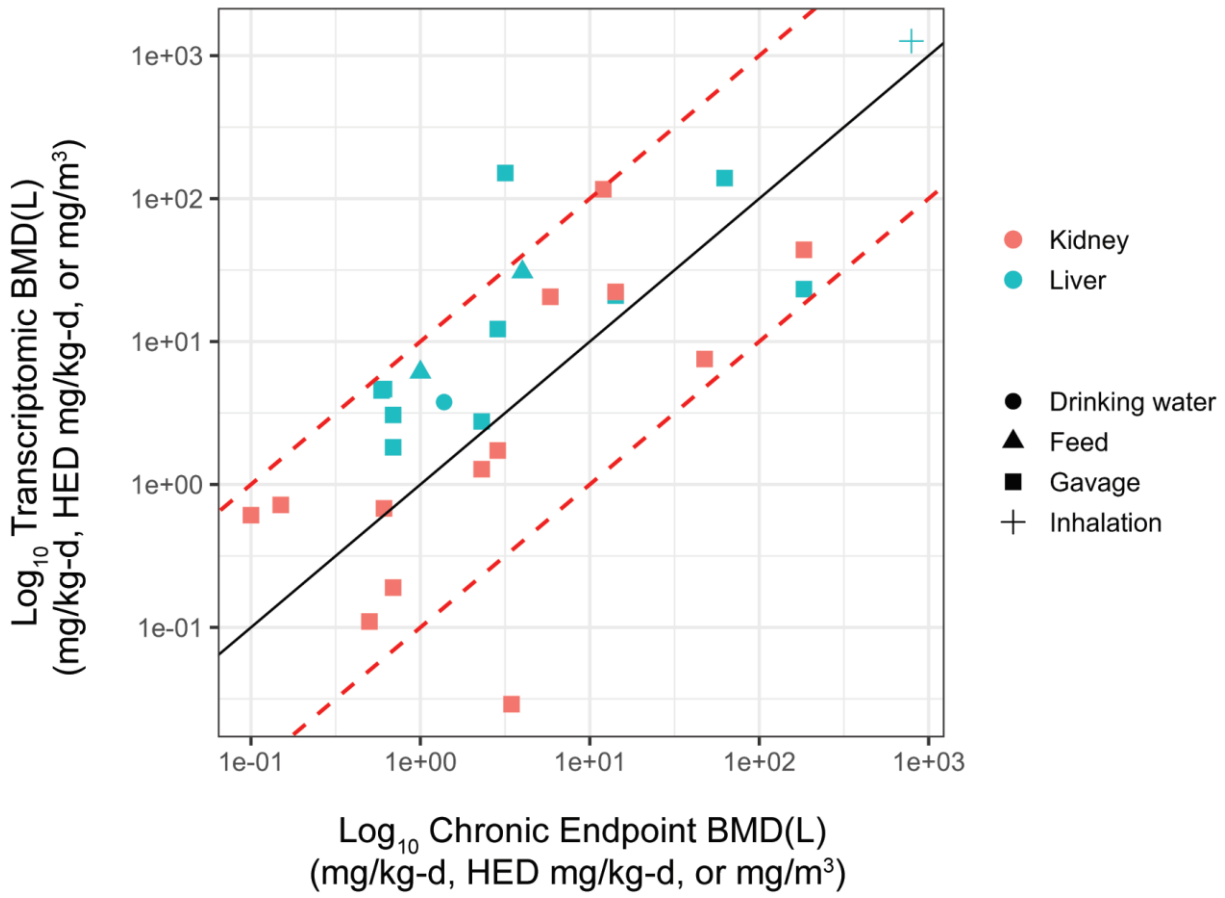


Figure 3-2. Concordance of transcriptomic and apical BMD(L)s for liver and kidney as sentinel tissues. The concordance between the transcriptomic BMD(L)s and chronic apical BMD(L)s is based on transcriptomic data from surrogate tissues (*i.e.,* liver or kidney) using data from Bianchi et al. 2021; Chepelev et al. 2017; Gwinn et al. 2020; Thomas et al. 2011; Zhou et al. 2017. The data is presented as BMDs unless BMDLs were the only values reported in the manuscript as for Bianchi et al. 2021 and Zhou et al. 2017. The apical BMD(L)s are all derived from two-year chronic bioassays, while the transcriptomic BMD(L)s are derived from studies using exposure durations from 1 - 90 days. The solid line indicates a perfect concordance at 1 and the dashed lines represent a 10-fold difference from the BMD(L).

3.8. IMPACTS OF TECHNOLOGY PLATFORM ON TRANSCRIPTIONAL AND APICAL DOSE CONCORDANCE

Large scale gene expression technologies have evolved over the past two decades from targeted, fluorescent intensity-based methods used by microarrays to both targeted and untargeted, count-based RNA-Seq approaches. RNA-Seq has the added benefit of measuring a larger dynamic range in gene expression measurements. In a recent analysis that compared microarrays versus RNA-Seq in a dose response experiment, Black and colleagues (2014) reported similar results. In that study, bromobenzene-induced fold changes in genes measured by microarray and RNA-Seq correlated well with each other across dose-levels despite differences in platform-specific normalization procedures (Spearman correlation coefficient 0.65 to 0.77). RNA-Seq tended to

identify more significant genes across all dose-levels when fold-change and statistical filters were applied to the dataset. The gene set based transcriptomic BMDs between microarray and RNA-Seq were similar at 78 – 80 and 57 – 76 mg/kg-day, respectively, with the ranges defined by the normalization approach employed (Black et al. 2014). Both technologies provided gene set-based transcriptomic BMD values similar to the apical BMD of 96.8 mg/kg-day (Black et al. 2014).

As RNA-Seq has become more cost effective, it has replaced microarray-based gene expression methods as the main platform for assessing genome-wide changes in gene expression. Among RNA-seq approaches, the BioSpyder TempO-Seq platform combines some of the beneficial attributes of RNA-Seq dynamic range with targeted gene detection inherent to microarrays (Yeakley et al. 2017). The short, equal length probes used by TempO-Seq circumvent issues of read length biases found in RNA-Seq studies while focusing sequencing reads on probes, thereby reducing assay cost (Yeakley et al. 2017). Several studies demonstrate the comparability between TempO-Seq, RNA-Seq, and microarray in detecting dose-dependent alterations in chemically induced gene expression changes making it a cost-effective platform for transcriptomic potency assessments. Bushel and colleagues showed that TempO-Seq performed similarly to microarray and RNA-Seq with respect to analysis of the SEQC transcriptomics data (Bushel et al. 2018).

The studies examining the relationship between transcriptional and apical BMD(L)s reviewed in this report employed different platforms (*i.e.*, microarrays, RNA-Seq or TempO-Seq) to evaluate transcriptional changes. While these studies also examined different chemicals using different exposure paradigms, the concordance between transcriptomic and apical BMD(L) were generally not influenced by the technology used to evaluate transcriptional changes (Figure 3-3).

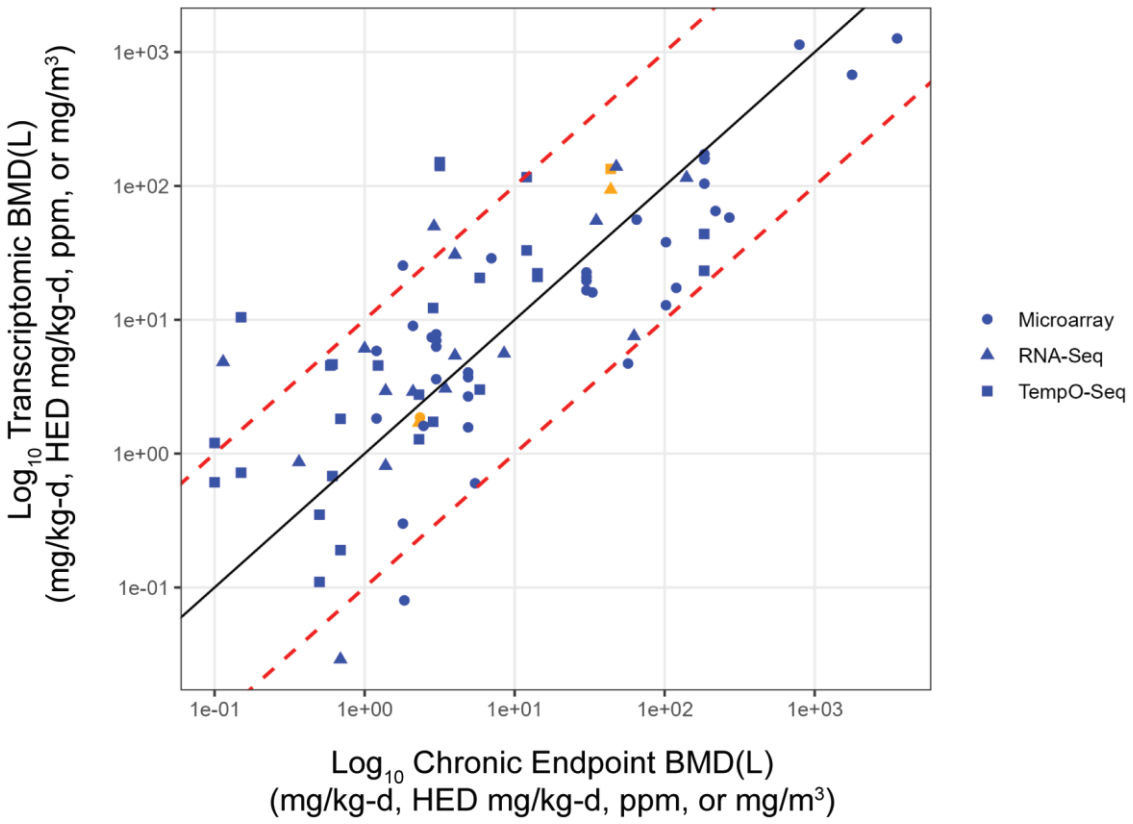


Figure 3-3. Concordance of transcriptomic and apical BMD(L)s across transcriptomic platforms. Data was obtained from (Andersen et al. 2010; Bercu et al. 2010; Bhat et al. 2013; Bianchi et al. 2021; Cannizzo et al. 2022; Chepelev et al. 2017, 2018; Dong et al. 2016; Gwinn et al. 2020; Hester et al. 2016; Jackson et al. 2014; Labib et al. 2017; LaRocca et al. 2020; Moffat et al. 2015; Recio et al. 2017; Thomas et al. 2007; Thomas et al. 2011; Thomas et al. 2013a; Thomas et al. 2013b; Zhou et al. 2017). The data is presented as BMDs unless BMDLs were the only values reported in the manuscript as for Bhat et al. 2013; Bianchi et al. 2021; and Zhou et al. 2017. While the apical BMD(L)s all are derived from two-year chronic bioassays, the transcriptomic BMD(L)s are derived from studies using exposure durations from 1 - 90 days. The solid line indicates a perfect concordance at 1, and the dashed lines represent a 10-fold difference from the BMD(L). Orange colored symbols represent paired data from experiments where the same samples underwent microarray and RNA-Seq or TempO-Seq and RNA-Seq analyses.

3.9. SUMMARY OF TRANSCRIPTIONAL AND APICAL DOSE CONCORDANCE

In studies spanning over a decade, researchers have evaluated the concordance of short-term transcriptomic BMD(L)s and apical BMD(L) for over a hundred chemicals. Some studies were hypothesis-based mechanistic studies [*e.g.*, (Dunnick et al. 2017; Geter et al. 2014; Moffat et al. 2015)], while others were more agnostic of the specific mechanisms and evaluated general concordance across larger, more diverse sets of chemicals [*e.g.*, (Gwinn et al. 2020; Johnson et al. 2020; Thomas et al. 2012)]. Known target tissues were often evaluated [*e.g.*, (Dunnick et al. 2017; Thomas et al. 2013b)], while others employed sentinel tissues such as the liver and kidney (Gwinn et al. 2020; Johnson et al. 2020). Exposure durations ranged from 1- to 90-days and the chemical,

biological, and toxicological space examined covers much of the known chemistries and toxicities of concern by EPA and other entities that assess environmental chemicals. Even though these studies were not coordinated, the comparison across studies indicates that transcriptomic PODs from short-term *in vivo* studies provide robust estimates of apical PODs from traditional chronic rodent toxicity studies. The $\log_{10}$ RMSD of 0.593 ($\log_{10}$ mg/kg-day) for the 1- to 90-day transcriptomic BMD values versus chronic apical BMD values in the combined data set was similar to the range of inter-study standard deviation estimates for the LOAELs for systemic toxicity in repeated dose studies, approximated as residual RMSE in log10-mg/kg-day units [0.45-0.56; (Pham et al. 2020)]. Overall, the results suggest that the error associated with the concordance between the transcriptomic BMD values versus non-cancer and cancer apical BMD values is approximately equivalent to the inter-study variability in the repeated dose toxicity study itself.

4. DEVELOPMENT OF TRANSCRIPTOMIC POINTS OF DEPARTURE FOR ETAP

4.1. OVERVIEW OF THE APPROACH TO DERIVE TRANSCRIPTOMIC POINTS OF DEPARTURE FOR ETAP

The approach used to derive the transcriptomic PODs for ETAP is largely based on the methodology outlined in the peer-reviewed report entitled National Toxicology Program Approach to Genomic Dose Response Modeling (NTP 2018). The NTP approach outlined in that report originated from a series of publications that developed and refined study designs for quantitative transcriptomic evaluation for dose response assessment and adapted dose response modeling methods used by the EPA for apical endpoints to transcriptomic data (Black et al. 2014; Rowlands et al. 2013; Thomas et al. 2007; Thomas et al. 2011; Thomas et al. 2012; Thomas et al. 2013b). The study design outlined in the NTP approach and used in the ETAP is a 5-day, repeated dose *in vivo* study in male and female rats with an extended dose response range at multiple dose levels. In the original NTP studies, transcriptomic measurements were performed on the liver and kidneys as sentinel tissues using the BioSpyder TempO-Seq rat S1500+ platform. However, for the 5-day, repeated dose studies for ETAP, transcriptional measurements will be performed on a larger number of tissues to increase the breadth of biological responses evaluated. The tissues will include kidney, liver, adrenal gland, brain, heart, lung, ovary (females), spleen, testis (males), thyroid, thymus, and uterus (females). The TempO-Seq rat S1500+ platform will be used in the ETAP as a pragmatic choice that provides a balance between a set of curated genes that can be cost-effectively employed across multiple tissues, doses, and chemicals and the need to cover important toxicological and disease processes (Mav et al. 2018).

The transcriptomic dose response modeling approach employed in the ETAP is aligned with the NTP report and follows four steps: 1) pre-modeling dataset evaluation to determine adequate signal; 2) pre-modeling probe filtering to remove those that are not responding to treatment; 3) dose response modeling of the individual probes, identifying the best-fit model, and deriving BMD(L) values; and 4) combining the individual probes into gene sets and summarizing the transcriptional BMD(L) values. Consistent with the NTP approach, the transcriptional dose response results from the gene set with the lowest median BMD value will be used to derive the POD for the ETAP (NTP 2018). Transcriptomic BMD values from the gene set with the lowest median BMD value following 5 days of exposure have been demonstrated to be concordant with non-cancer and cancer phenotypic responses in subchronic and chronic toxicity studies (see review in Section 3). The coordinated transcriptional changes used to identify the POD do not necessarily discriminate between specific hazards, adaptive or adverse effects, nor are they used to infer a mechanism or mode of action.

Rather, the transcriptomic POD is used to define the experimentally determined dose at which there were no coordinated transcriptional changes that could indicate a potential toxicity of concern.

4.2. ANALYSIS TO SELECT STUDY DESIGN AND TRANSCRIPTOMIC PLATFORM SPECIFIC DOSE RESPONSE MODELING PARAMETERS

The NTP Approach to Genomic Dose Response Modeling report provided general recommendations for the selection of settings and parameters for each step in the dose response modeling process; however, some of the recommendations were acknowledged to be platform specific or provided with minimal data supporting them. The NTP report suggested evaluating various settings and parameter choices to identify an optimal combination that increases detection of true signal, minimizes false signal, and maximizes reproducibility (NTP 2018). To address this suggestion, a comprehensive analysis was undertaken to identify and support the choices and parameters used in each step of the transcriptomic dose response modeling process. Subsets of the data from two NTP datasets were specifically used to address the three goals outlined in the NTP approach (Fig. 4-1; orange boxes). From the first NTP dataset (Gwinn et al. 2020), transcriptomic dose response data for 14 chemicals with data from chronic rodent bioassays were used to evaluate settings and parameters for each component of the dose response modeling process with respect to dose concordance of transcriptional and apical responses (*i.e.*, increase detection of true signal). From the second NTP dataset[26], transcriptomic data for three chemicals each with multiple independent replicates were used to evaluate settings and parameters for each component of the dose response modeling process with respect to inter-study reproducibility (*i.e.*, maximize reproducibility). Lastly, combined vehicle control data from both studies were used to evaluate settings and parameters for each component of the dose response modeling process with respect to the family-wise error rate (*i.e.*, minimizing false signal). The two publications used the 5-day, repeated dose *in vivo* rat study design and BioSpyder TempO-Seq rat S1500+ platform. Therefore, conclusions regarding optimal settings and parameter choices are directly applicable to the ETAP studies.

[26] The second NTP dataset is available at: https://doi.org/10.22427/NTP-DATA-002-00099-0001-000-1

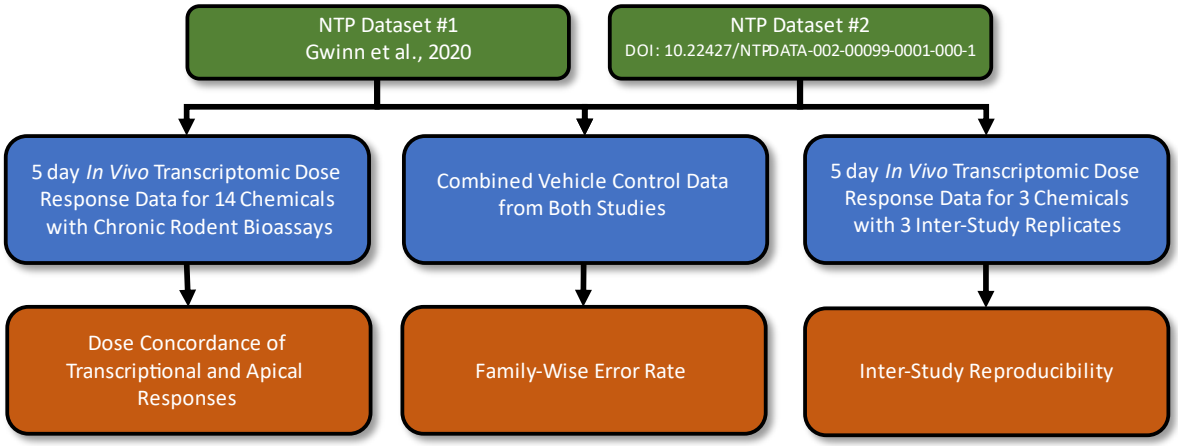


Figure 4-1. Overview of the analysis performed to select study design and transcriptomic platform specific dose response modeling parameters. Subsets of data from two NTP datasets (Gwinn et al. 2020; https://doi.org/10.22427/NTP-DATA-002-00099-0001-000-1) were used in the analysis. Addressing the three goals outlined in the NTP report (*i.e.,* increasing detection of true signal, minimizing false signal, and maximizing reproducibility) are denoted by the orange boxes.

4.2.1. DOSE CONCORDANCE OF TRANSCRIPTIONAL AND APICAL RESPONSES

4.2.1.1. Overview

To evaluate which dose response modeling parameters provided the best dose concordance between transcriptional and apical responses, the transcriptomic data for the individual chemicals were analyzed using 48 different combinations of selected pre-modeling probe filtering, dose response modeling, and gene set summarization parameters. For each parameter combination, the transcriptomic BMD based on the gene set with the lowest median BMD value was then compared with the minimum of the non-cancer and cancer apical BMD values from the chronic toxicity study. The combinations of parameters were rank ordered based on the RMSD. The parameter combination with the lowest RMSD was selected as the optimal combination for use in the ETAP.

4.2.1.2. Identification of Chronic Apical BMD Values

In the original publication, a subset of 17 out of 19 chemicals had corresponding two-year chronic rodent bioassays in the male rat (Gwinn et al. 2020). Three of these substances, tetrabromobisphenol A, ginseng, and milk thistle extract, did not result in statistically significant apical effects in male rats after two-years of exposure. Since publication of the study, histopathological results from a two-year chronic rodent bioassay for tris(2-chloroisopropyl)

phosphate were released,[27] and in 2021 a NTP technical report for a two-year chronic rodent bioassay on di(2-ethylhexyl) phthalate was published (NTP 2021). The Gwinn et al. study relied on an earlier NTP technical report for di(2-ethylhexyl) phthalate that was published in 1982. The total number of chemicals from Gwinn et al. study with a two-year chronic rodent bioassay in the male rat in which a statistically significant apical effect was reported is 14. The minimum of the non-cancer and cancer apical BMD values from the chronic study are provided in Table 4-1. Details on the calculation of the apical BMD values are provided in the Appendix (Section 6.1).

Table 4-1. Adverse responses based on the minimum BMD values among non-cancer and cancer histopathological endpoints for 14 chemicals with chronic two-year rodent bioassays.

Chemical	Rat Strain or Stock	Route	Endpoint	Minimum Route Adjusted Apical BMD (BMDL) (mg/kg-day)
Acrylamide	F344/N	Drinking water	Peripheral nerve (sciatic) axon degeneration	0.61 (0.43)
Bromodichloroacetic acid	F344/NTac	Drinking water	Bone marrow angiectasis	2.30 (1.87)
Coumarin	F344/N	Gavage	Liver necrosis	5.85 (4.85)
Pentabromodiphenyl ether mixture (DE71)	Wistar Han	Gavage	Liver hepatocyte hypertrophy	0.15 (0.11)
Di(2-ethylhexyl) phthalate[a]	Sprague Dawley	Feed	Pancreatic acinar adenoma or carcinoma	31.2 (20.3)
Ethinyl estradiol	NCTR SD (F1C)	Feed	Mammary gland alveolar hyperplasia	0.00069 (0.00047)
Furan	F344/N	Gavage	Liver cholangiofibrosis	0.10 (0.09)
Hexachlorobenzene	Sprague Dawley	Feed	Chronic nephrosis	0.59 (0.35)
Methyl eugenol	F344/N	Gavage	Liver hepatocellular adenoma or carcinoma	12 (9.9)
Perfluorooctanoic acid	Sprague Dawley	Feed	Liver hepatocyte hypertrophy	0.50 (0.41)
Pulegone	F344/N	Gavage	Nose olfactory epithelium degeneration	9.5 (7.5)
3,3',4,4'-Tetrachloroazobenzene	Sprague Dawley	Gavage	Forestomach epithelium hyperplasia	2.9 (2.1)
α,β-Thujone	F344/N	Gavage	Kidney mineralization	3.2 (2.3)

[27] NTP histopathological data for tris(2-chloropropyl) phosphate: https://ntp.niehs.nih.gov/go/TS-m20263

Table 4-1. Adverse responses based on the minimum BMD values among non-cancer and cancer histopathological endpoints for 14 chemicals with chronic two-year rodent bioassays.				
Tris(2-chloroisopropyl) phosphate[b]	Sprague Dawley	Feed	Lung focal granulomatous inflammation	240 (141)
[a]The chronic two-year bioassay results for BMD modeling were obtained from the NTP Technical Report 601. [b]The chronic two-year bioassay results for BMD modeling were obtained from the NTP histopathological tables posted for peer review: https://ntp.niehs.nih.gov/go/TS-m20263.				

4.2.1.3. Transcriptomic BMD Modeling Calculations

The transcriptomic dose response modeling process can be broken down into four main steps: 1) pre-modeling dataset evaluation to determine adequate signal; 2) pre-modeling probe filtering to remove those that are not responding to treatment; 3) dose response modeling of the individual probes, identifying the best-fit model, and deriving BMD(L) values; and 4) combining the individual probes into gene sets and summarizing the transcriptional BMD(L) values. At each step in the process, different choices and parameter values are employed for a variety of statistical, biological, and practical reasons including ensuring adequate fit of the dose response model; removing noisy genes or probes; and ensuring sufficient transcriptional responses at the gene set level. As noted in the NTP report, some of the settings and parameter values were acknowledged to be study design or transcriptomic platform specific, while others are not typically study design or platform dependent.

The raw sequencing reads (FASTQ files) for the 14 chemicals with adverse apical outcomes in chronic rodent bioassays were obtained from the NTP. The methods for aligning, normalizing, and quality control of the sequencing data are outlined in the Appendix (Section 6.2). Following quality control, a pre-modeling dataset evaluation was performed on each treatment group using an analysis of variance (ANOVA) with a cut-off of at least one probe showing statistical significance at a Benjamini and Hochberg False Discovery Rate (FDR) corrected p-value < 0.05 (NTP 2018)(See Section 4.2.3 for evaluation of this step). If the treatment group passed the pre-modeling evaluation, transcriptomic dose response modeling was performed. In the transcriptomic dose response modeling, selected BMD modeling settings and parameters were varied to evaluate the impact on the dose concordance between transcriptomic and apical responses. A total of 48 different settings and parameter combinations were evaluated.

For the pre-modeling probe filtering step, the NTP genomics report recommended a combination of statistical significance based on a William's Trend test combined with a minimum effect size (*i.e.*, fold-change relative to control)(NTP 2018). However, the values associated with these settings may be study design or transcriptomic platform specific. In this analysis, William's Trend

test p-value cut-offs of 0.05 and 0.1 were evaluated together with minimum absolute fold-change cut-offs of 1.5- and 2.0-fold (Table 4-2).

Table 4-2. Transcriptomic BMD modeling settings and parameters that were fixed or varied in the dose concordance analysis.

BMD Modeling Step / Parameter	Fixed/Varied	Value(s)
Pre-Modeling Probe Filtering		
Maximum William's Trend Test p-Value[a]	Varied	0.05; 0.1
Minimum Absolute Fold-Change	Varied	1.5; 2.0
Benchmark Dose Modeling		
Benchmark Response (BMR)	Fixed	1.349
Model Selection	Fixed	Lowest AIC (Akaike information criterion)
Minimum Fit p-Value	Fixed	0.1
Maximum BMD Filter	Fixed	Highest Dose
Minimum Hill Model 'k' Parameter Filter	Fixed	One Third Lowest Dose
Maximum BMD Uncertainty Filter	Varied	BMD/BMDL>20; BMDU/BMDL>40
Gene Set Summarization		
Gene Sets	Fixed	GO Biological Process
Summary Value	Fixed	Median BMD/BMDL
Minimum Number of Genes	Varied	3; 5
Minimum Gene Set Coverage	Varied	0%; 3%; 5%

[a]The William's trend test p-value is not adjusted for multiple comparisons.

For the BMD modeling, many of the choices and parameters were fixed given they are not inherently study design or transcriptomic platform dependent. Model fitting was performed on each probe. Linear, second-degree polynomial, power, Hill, second degree exponential, third degree exponential, fourth degree exponential, and fifth degree exponential models were fit to the dose response curves assuming constant variance. The exponent for the power model was restricted (>=1). The model with the lowest Akaike information criterion (AIC) was selected as the best-fit model except in cases where the "k" parameter for the Hill model is less than one-third the lowest dose. In these cases, where the "k" parameter for the Hill model was out of bounds, the Hill model was excluded from the final selection (Rowlands et al. 2013; Thomas et al. 2013b). The Benchmark Response (BMR) was set to 1.349 * standard deviation of replicate vehicle control samples (Thomas et al. 2007). The BMR is different from the recommendation in the NTP genomics report, which listed 1 standard deviation (NTP 2018). Based on EPA guidance, the 1 standard deviation for continuous data is equivalent to a 10% increase in risk for normally distributed effects when the direction of the effects is known (EPA 2012). However, for most gene expression changes, the direction is not known *a priori*. To provide an equivalent 10% increase in risk, a BMR of 1.349 * standard deviation is

required (Thomas et al. 2007). Probes with a BMD greater than the highest dose or a goodness-of-fit p-value less than 0.1 were removed from the analysis. Apart from the fixed BMD modeling parameters, the NTP genomics report recommended removing probes with a high uncertainty in the BMD by applying a BMDU/BMDL filter >40 (NTP 2018). However, the noise associated with BMD values may be study design and transcriptomic platform specific. As a result, two different BMD uncertainty filters (BMD/BMDL>20 and BMDU/BMDL>40) were evaluated (Table 4-2).

For gene set summarization, the GO biological processes were among the gene sets recommended in the NTP approach (NTP 2018). The median BMD and BMDL values were also recommended for summarizing the gene set level potencies (NTP 2018). However, the minimum number of genes in the gene set and the minimum gene set coverage may be transcriptomic platform dependent, especially given the measurement of a smaller number of genes using the S1500+ assay. In this analysis, different cut-offs for the minimum number of genes (3 and 5) and minimum gene set coverage (0%, 3%, and 5%) were evaluated (Table 4-2).

4.2.1.4. Evaluation of Dose Concordance for Transcriptional and Apical Responses

For each gene expression dataset, BMD values for each GO biological process class were calculated for each tissue and each of the 48 combinations of pre-modeling probe filtering, BMD modeling, and gene set summarization parameters. The lowest median BMD value among GO biological processes in either tissue (liver and kidney) was used as the transcriptomic BMD. The $\log_{10}$-transformed transcriptomic BMD values were then compared with the minima of the $\log_{10}$-transformed chronic non-cancer and cancer apical BMD values using Pearson's correlation coefficient and RMSD. For chemicals with replicate transcriptomic studies, the transcriptomic BMD values within each study were averaged together to derive a single transcriptomic BMD estimate for comparison to the apical BMD. The RMSD was calculated as follows:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - X_i)^2}{N}} \quad (1)$$

Where X_i is the $\log_{10}$ transcriptomic BMD value for the i^{th} chemical, Y_i is the minimum of the $\log_{10}$ chronic non-cancer and cancer apical BMD values for the i^{th} chemical, and N is the total number of chemicals. Combinations that failed to derive a transcriptomic BMD (*i.e.*, that had no GO biological process class passing all filters in either tissue) for any of the 14 chemicals were removed. The combinations of parameters were rank ordered based on RMSD.

Across all combinations of parameters that successfully derived a transcriptomic BMD for all 14 chemicals, the Pearson correlation coefficient ranged from 0.804 to 0.917, while the RMSD ranged from 0.567 to 0.958 ($\log_{10}$ mg/kg-day). The top five combinations of parameters based on the RMSD are provided in Table 4-3. In each of the top five combinations, an absolute fold-change >1.5, BMD/BMDL < 20, and minimum of 3 genes per GO class were consistently represented. Only the

William's p-value and percentage of genes in the set varied in their representation in the top five ranked combinations.

Table 4-3. Top five combinations of pre-modeling probe filter, BMD modeling, and gene set summarization parameters based on RMSD.

Rank	Pre-Modeling Probe Filtering, BMD Modeling, and Gene Set Summarization Parameter Combination[a]	Pearson Correlation Coefficient	RMSD (log_{10} mg/kg-day)
1	Williams p < 0.05; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 0%	0.910	0.567
2	Williams p < 0.1; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 0%	0.907	0.571
3	Williams p < 0.1; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 3%	0.905	0.578
4	Williams p < 0.1; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 5%	0.906	0.581
5	Williams p < 0.05; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 3%	0.905	0.593

[a]FC, fold-change.

The best overall combination of parameters based on the minimum RMSD of the transcriptomic versus apical BMD values includes pre-modeling probe filtering criteria of |Fold-Change| > 1.5 with a William's trend test p-value of <0.05 and a post-modeling filter to remove probes with BMD/BMDL ratio > 20. When summarizing the results for the rat S1500+ assay based on GO biological process class, the best combination of parameters had a minimum of 3 genes with a valid BMD as the cutoff, with no minimum requirement on percent coverage. Using the recommended parameter combinations, the Pearson correlation coefficient and RMSD of the transcriptomic versus chronic apical BMD values were 0.910 and 0.567, respectively (Figure 4-2). The median absolute ratio of the transcriptomic BMD and chronic non-cancer apical BMD values was 3.2 $\pm$ 1.9 (MAD). The maximum absolute fold-difference was 7.87. The RMSD value for the best parameter combination was slightly higher than that reported by Johnson and colleagues when comparing transcriptomic and apical POD values for 29-day toxicity studies [0.54 and 0.48 for consistent and inconsistent dose levels, respectively; (Johnson et al. 2020)]. For further comparison, the RMSD value is similar to the range of inter-study standard deviation estimates for the LOAELs for systemic toxicity in repeated dose studies, approximated as residual RMSE in log_{10}-mg/kg-day units [0.45-0.56; (Pham et al. 2020)]. The results suggest that the error associated with the concordance between the transcriptomic BMD values versus non-cancer and cancer apical BMD values for the best parameter combination is approximately equivalent to the inter-study variability in the repeated dose toxicity study itself.

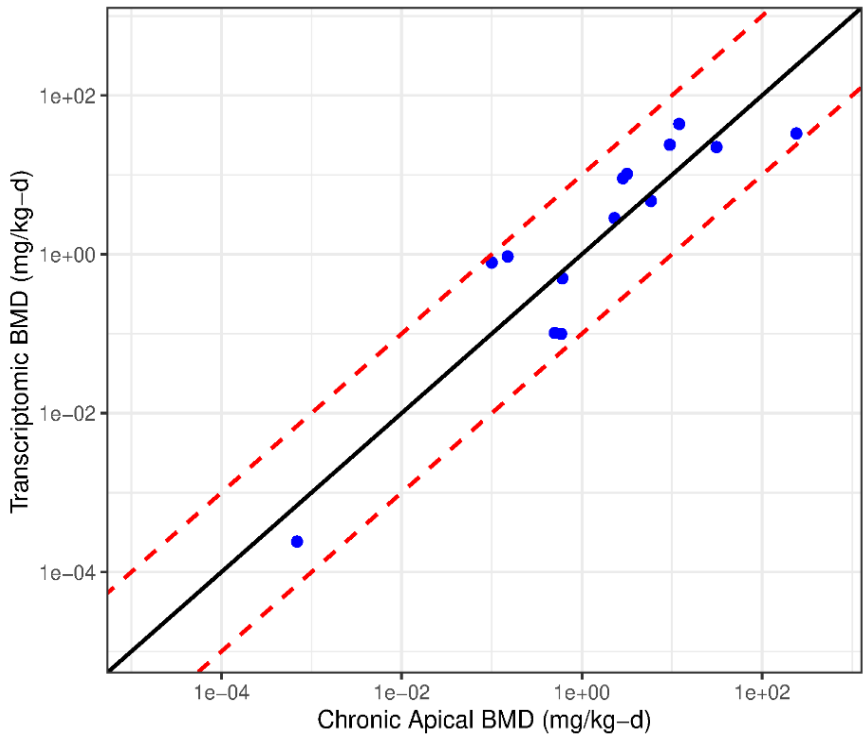


Figure 4-2. Scatter plot of $\log_{10}$ transcriptomic BMD versus the minimum of the chronic non-cancer and cancer apical $\log_{10}$ BMD values for the top ranked combination of pre-modeling probe filter, BMD modeling, and gene set summarization parameters (Table 4-3). The black line is 1:1 concordance. The red lines are ± 10-fold. Values below the black line indicate the transcriptomic BMD value is less than the chronic apical BMD value.

Concordance between the transcriptomic and chronic apical BMDL values were also compared. Using the best parameter combination, the Pearson correlation coefficient and RMSD of the transcriptomic versus chronic apical BMDL values were 0.908 and 0.694, respectively (Figure 4-3). The median absolute fold-difference between the transcriptomic BMDL and chronic apical BMDL values was 2.8 ± 1.6 (MAD). Notably, the majority of transcriptomic BMDL values are lower than the chronic apical BMDL values. This contrasts with the BMD values where the transcriptomic BMD values were approximately equally distributed above and below the unity line. This suggests that the confidence intervals for the transcriptomic BMD values are slightly wider than those for the apical BMD values leading to more conservative PODs used to derive reference values.

4.2.2. EVALUATION OF INTER-STUDY REPRODUCIBILITY

4.2.2.1. Overview

To evaluate the inter-study reproducibility of the transcriptomic BMD and BMDL values, the transcriptomic data from three independently replicated chemicals were analyzed. The replicated chemicals included furan (n = 3), perfluorooctanoic acid (n = 3), and bromodichloroacetic acid (n = 3). The replicate studies were run at the same contract lab over the course of several years and

required the preparation of new dosing solutions for each study using the same supplier and lot for each chemical.

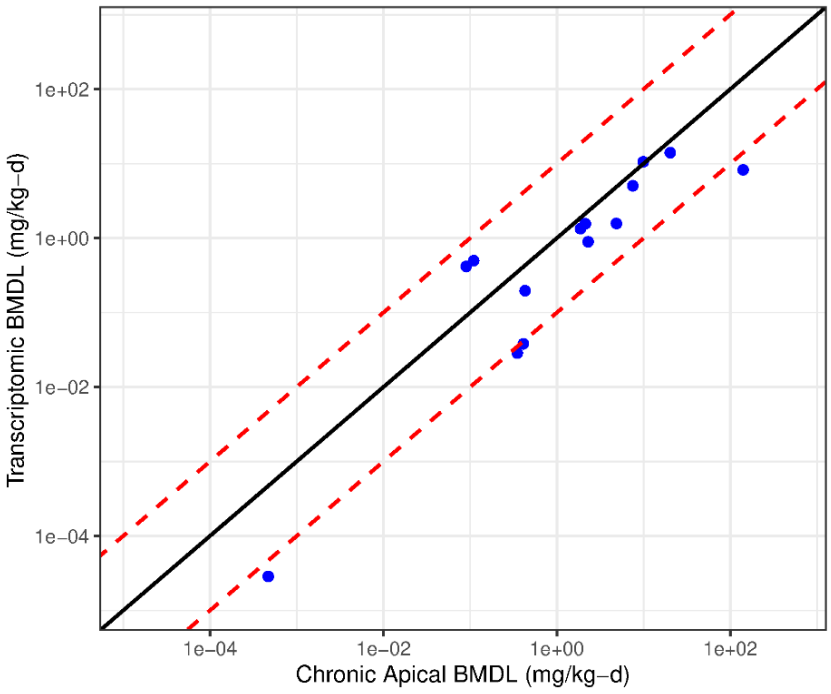


Figure 4-3. Scatter plot of $\log_{10}$ transcriptomic BMDL versus the minimum of the chronic non-cancer and cancer apical $\log_{10}$ BMDL values for the top ranked combination of pre-modeling probe filter, BMD modeling, and gene set summarization parameters (Table 4-3). The black line is 1:1 concordance. The red lines are ± 10-fold. Values below the black line indicate the transcriptomic BMDL value is less than the chronic apical BMDL value.

4.2.2.2. Calculation of Inter-Study Reproducibility

The raw sequencing reads (FASTQ files) for the three independently replicated chemicals with chronic rodent bioassays were obtained from the NTP. The detailed methods for aligning, normalizing, and quality control of the sequencing data are outlined in the Appendix (Section 6.2). Following quality control, the dose response series for each of the replicates were analyzed using the complete transcriptomic dose response analysis process including: 1) pre-modeling dataset evaluation to determine adequate signal; 2) pre-modeling probe filtering to remove those that are not responding to treatment; 3) dose response modeling of the individual probes, identifying the best-fit model, and deriving BMD(L) values; and 4) combining the individual probes into gene sets and summarizing the transcriptional BMD(L) values. For the pre-modeling dataset evaluation, an ANOVA evaluation with an FDR < 0.05 cut-off for 1 or more probes was used. For the rest of the dose response modeling process, the top 5 combinations of pre-modeling probe filtering, BMD modeling, and gene set summarization parameters in Table 4-3 were evaluated. The inter-study reproducibility

was calculated using the estimated standard deviation (SD) of the transcriptomic BMD and BMDL values. The SD was estimated as follows:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - X_i)^2}{2N}} \quad (2)$$

Where X_i and Y_i are the $\log_{10}$ transcriptomic BMD or BMDL values for the i^{th} pair of independently replicated studies for the same chemicals, and N is the total number of pairs of independently replicated studies.

The results suggest that the parameter combinations which produced transcriptomic BMDs that were most concordant with the apical BMDs also produce transcriptional BMDs that are highly reproducible based on the independent replicate studies for the three chemicals (Table 4-4). Furthermore, all five of the parameter combinations result in highly similar SD values for both the transcriptional BMD and BMDL values. Therefore, we chose the parameter combination with the overall lowest RMSD for transcriptional versus apical BMD values, as the differences in BMD/BMDL between the top combinations were negligible.

Table 4-4. Inter-study variability of the median BMD and BMDL values for the top five combinations of pre-modeling probe filter, BMD modeling, and gene set summarization parameters.

Rank	Pre-Modeling Probe Filtering, BMD Modeling, and Gene Set Summarization Parameter Combination[a]	Log_{10} BMD SD (log_{10} mg/kg-day)	Log_{10} BMDL SD (log_{10} mg/kg-day)
1	Williams p < 0.05; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 0%	0.242	0.295
2	Williams p < 0.1; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 0%	0.247	0.292
3	Williams p < 0.1; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 3%	0.245	0.290
4	Williams p < 0.1; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 5%	0.241	0.289
5	Williams p < 0.05; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 3%	0.242	0.289

[a]FC = fold change.

4.2.3. EVALUATION OF FAMILY-WISE ERROR RATE

4.2.3.1. Overview

To estimate the family-wise error rate, individual samples from the corn oil vehicle control groups were randomly distributed into 1,000 sham dose response series for each tissue. The sham dose response series were then analyzed to estimate the family-wise error rate for the pre-modeling dataset evaluation step as well as the complete transcriptomic dose response analysis process.

4.2.3.2. Pre-Modeling Dataset Evaluation of Sham Dose Response Series

According to EPA benchmark dose guidance, the dataset being modeled should have a statistically or biologically significant dose-related trend (EPA 2012). For transcriptomic dose response data, the NTP report recommended performing an ANOVA prior to dose response modeling with a cut-off of at least one gene or probe showing statistical significance at a Benjamini-Hochberg FDR corrected p-value < 0.05 (NTP 2018). To evaluate the fitness of this recommendation for the specific study design and transcriptomics platform employed in the ETAP, the raw sequencing reads (FASTQ files) for the corn oil vehicle control groups were obtained from NTP. The methods for aligning, normalizing, and quality control of the sequencing data are outlined in the Appendix (Section 6.2). A subset of 53 liver and kidney samples from the corn oil vehicle control groups (14 studies that used corn oil as the matched vehicle control X 4 animals per study minus low quality and outlier samples) were randomly distributed into 1,000 sham dose response series for each tissue. Each sham dose response series consisted of nine groups (one control group and eight mock positive dose groups) with four samples per group. To represent realistic dose ranges, an equal fraction of sham dose response series was assigned the 8 lowest doses for each of the 14 chemical regimens tested. The sham dose response series were analyzed using ANOVA and a range of FDR corrected p-values. The estimated family-wise error rate was computed based on the number of the 1,000 sham dose response series with at least one probe passing the ANOVA with FDR correction (Fig. 4-4). The results suggest that a pre-modeling ANOVA evaluation with an FDR < 0.05 cut-off for 1 or more probes results in an estimated family-wise error rate of 0.046 in both the liver and kidney. The estimated family-wise error rate for the sham dose response series approximates the target FDR for the analysis.

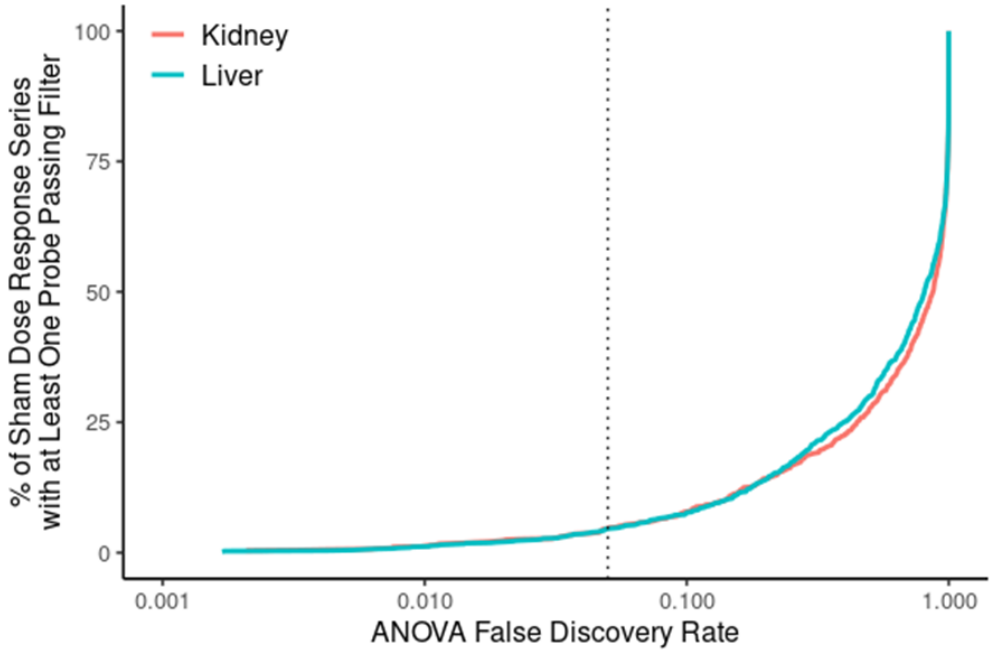


Figure 4-4. Percentage of sham dose response datasets with at least one probe showing statistical significance based on an ANOVA with a p-value of varying FDR correction. The sham dose response datasets were created by randomly distributing the individual corn oil vehicle control samples into 1,000 dose response treatment groups. The blue line are the percentages associated with the liver tissue. The orange line are the percentages associated with the kidney tissue. The dotted line delineates an FDR = 0.05.

4.2.3.3. Complete Transcriptomic Dose Response Analysis of the Sham Dose Response Series

Although characterizing the family-wise error rate of the initial pre-modeling dataset evaluation step provides an understanding of the number of potential datasets that may make it into the dose response modeling process erroneously, the results may not reflect the overall family-wise error rate of the complete transcriptomic dose response analysis process. To estimate the overall family-wise error rate for identifying a gene set-level BMD with the specific study design and transcriptomics platform employed in the ETAP, the same set of 1,000 sham dose response series were analyzed using the complete transcriptomic dose response analysis process including: 1) pre-modeling dataset evaluation to determine adequate signal; 2) pre-modeling probe filtering to remove those that are not responding to treatment; 3) dose response modeling of the individual probes, identifying the best-fit model, and deriving a BMD value together with its lower confidence bound (*i.e.*, BMD and BMDL); and 4) combining the individual probes into gene sets and summarizing the transcriptional BMD and BMDL values. For the pre-modeling dataset evaluation, an ANOVA evaluation with an FDR < 0.05 cut-off for 1 or more probes was used. For the rest of the dose response modeling process, the top 5 combinations of pre-modeling probe filter, BMD modeling, and gene set summarization parameters in Table 4-3 were evaluated. A false positive was counted when a sham

dose response series had at least one GO biological process class with a valid BMD and BMDL. The results suggest that combining the pre-modeling dataset evaluation step with the transcriptomic dose response modeling process significantly reduced the family-wise error rate from 0.046 to less than 0.01. All of the top five combinations of pre-modeling probe filter, BMD modeling, and gene set summarization parameters have an overall family-wise error rate of less than 0.01 with the highest ranked combination of parameters with a family-wise error rate of 0.006 (Table 4-5).

Table 4-5. Overall family-wise error rate of the top five combinations of pre-modeling probe filter, BMD modeling, and gene set summarization parameters.

Rank	Pre-Modeling Probe Filtering, BMD Modeling, and Gene Set Summarization Parameter Combination[a]	Overall Family-Wise Error Rate
1	Williams p < 0.05; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 0%	0.006
2	Williams p < 0.1; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 0%	0.009
3	Williams p < 0.1; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 3%	0.002
4	Williams p < 0.1; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 5%	0.002
5	Williams p < 0.05; FC > 1.5; BMD/BMDL < 20; min 3 genes; min 3%	0.001

[a]FC, fold change.

4.3. COMPARISON OF TRANSCRIPTIONAL AND APICAL DOSE CONCORDANCE IN THE CONTEXT OF INTER-STUDY VARIABILITY

The evaluation of the concordance between transcriptional BMD values from the short-term *in vivo* studies and the apical BMD values from the chronic rodent bioassays are confounded by the inter-study variability in both dimensions. Estimating and explicitly considering this variability is important for interpreting the concordance metrics (*e.g.*, RMSD) and the level of confidence in the application of the ETAP. To provide this context, the concordance MSD of the top combination of pre-modeling probe filter, BMD modeling, and gene set summarization parameters was compared with an estimate of the lower bound of the expected MSD given inter-study variances.

4.3.1. DERIVATION OF MSD LOWER BOUND

Let X_c be the observed transcriptomic BMD ($\log_{10}$ mg/kg-day) and Y_c the observed apical BMD ($\log_{10}$ mg/kg-day) for chemical c, where study design was standardized across chemicals. Following the work of Pham and colleagues (Pham et al. 2020), the apical BMD values were assumed to be random variables with means dependent on chemical and study design (but note study design is standardized across chemicals in this study) and constant variance after accounting for chemical and study design (*i.e.*, common variance across chemicals). In the absence of evidence to the contrary, the same was assumed for the transcriptomic BMD values. That is, it was assumed that $E[X_c] = \mu_X(c)$ and $E[Y_c] = \mu_Y(c)$, where $\mu_X(c)$ and $\mu_Y(c)$ are the mean transcriptomic and apical BMD values for

chemical *c*, respectively; and $Var[X_c] = \sigma_X^2$ and $Var[Y_c] = \sigma_Y^2$ are the inter-study, within-chemical variances for transcriptomic and apical BMD values, respectively.

Let $Z_c = X_c - Y_c$ be the difference between observed transcriptomic and apical BMD values for chemical *c*. Then $E[Z_c] = \mu_Z = \mu_X(c) - \mu_Y(c)$ (note that the difference in BMD means was assumed to be constant across chemicals) and $Var(Z_c) = \sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$ (note that X_c and Y_c are conditionally independent given chemical means, so no covariance term is needed).

The MSD concordance statistic between X_c and Y_c for *n* chemicals is an unbiased estimator of $E\left[Z_c^2\right]$:

$$MSD = \sum_{c=1}^{n} \frac{(x_c - y_c)^2}{n} = \sum_{c=1}^{n} \frac{z_c^2}{n} \quad (3)$$

That is, $E[MSD] = E\left[Z_c^2\right]$. The variance of Z_c can be decomposed as follows:

$$Var(Z_c) = E[Z_c^2] - \mu_Z^2 \quad (4)$$

Rearranging:

$$E[Z_c^2] = Var(Z_c) + \mu_Z^2 \quad (5)$$

It follows that $E\left[Z_c^2\right] = Var(Z_c)$ when $\mu_Z = 0$ (*i.e.*, when the mean values of X_c and Y_c are the same for each chemical), and $E\left[Z_c^2\right] > Var(Z_c)$ when $\mu_Z \neq 0$ (*i.e.*, when the expected values of X_c and Y_c differ across chemicals). Thus,

$$E[Z_c^2] \geq Var(Z_c) \quad (6)$$

Substituting $E\left[Z_c^2\right] = E[MSD]$:

$$E[MSD] \geq Var(Z_c) \quad (7)$$

$$E[MSD] \geq \sigma_X^2 + \sigma_Y^2 \quad (8)$$

Thus, the lower bound of expected MSD is the sum of the transcriptomic and apical BMD variances.

4.3.2. ESTIMATES OF INTER-STUDY VARIANCES AND LOWER BOUND OF EXPECTED CONCORDANCE MSD

Inter-study replicates were used to estimate the transcriptomic BMD variance, σ_X^2. For replicates *i* and *j* of chemical *c*, $E\left[X_{c,i} - X_{c,j}\right] = 0$ and $Var\left(X_{c,i} - X_{c,j}\right) = 2\sigma_X^2$ from Section 4.3.1 above. Let *k* be the number of chemicals with replicate transcriptomic BMD estimates, let r_c be the number of observed replicates for chemical *c*, and let $I_c = \{1, 2, \ldots, r_c\}$. An unbiased estimator of σ_X^2 is:

$$\hat{\sigma}_X^2 = \left(2\sum_{c=1}^{k}\binom{r_c}{2}\right)^{-1} \sum_{c=1}^{k} \sum_{i \in I_c} \sum_{j \in I_c; j>i} (x_{c,i} - y_{c,j})^2 \quad (9)$$

That is, σ_X^2 is estimated as one half the mean squared difference in transcriptomic BMD values between unique pairs of replicates for each chemical. Across the dose response modeling parameter combinations considered, the range of transcriptomic BMD variance estimates was 0.015-0.352 (or 0.123-0.594 for the standard deviation). The range of chronic apical LOAEL unbiased variance estimates from Pham et al. was used to approximate the apical BMD variance, σ_Y^2: 0.252-0.265 (or 0.502-0.515 for the standard deviation) (Pham et al. 2020). Combining these two variance estimates, the estimate of the lower bound for expected MSD was 0.267-0.617.

4.3.3. MSD OF THE TOP COMBINATION OF TRANSCRIPTOMIC DOSE RESPONSE MODELING PARAMETERS COMPARED WITH LOWER BOUND OF THE EXPECTED CONCORDANCE MSD GIVEN INTER-STUDY VARIANCES

The MSD of the top combination of pre-modeling probe filter, BMD modeling, and gene set summarization parameters using mean $\log_{10}$ BMD values for chemicals that had replicates was $0.567^2=0.321$ $(\log_{10}$ mg/kg-day$)^2$. However, using mean BMD values for only some chemicals violates the assumption of equal variance across chemicals used to derive the lower bound of expected MSD. For comparison with the lower bound estimate, the concordance MSD for the top model was computed using all combinations of single replicates per chemical, and the minimum and maximum of these point estimates were 0.285 and 0.386, respectively. Thus, the full range of MSD values computed using the single chemical replicates falls within the range of lower bound estimates for expected concordance MSD of 0.267 - 0.617 when considering the inter-study variances for both the transcriptomic studies and repeated dose studies examining systemic effects. The results suggest that the error associated with the concordance between the transcriptomic BMD values versus non-cancer and cancer apical BMD values is approximately equivalent to the combined inter-study variability associated with the 5-day transcriptomic study and the chronic rodent bioassay.

5. SUMMARY AND CONCLUSIONS

Regulatory agencies, such as EPA, face substantial challenges evaluating and managing the human health risks of the thousands of chemicals and mixtures on their respective inventories. Relatively few of the chemicals have traditional toxicity data and fewer have human health assessments that could be used to inform regulatory decisions. Technologies, such as transcriptomics, have potential to more rapidly evaluate potential human health effects and play a role in filling the data gaps in toxicity testing and human health assessment. In recognition of the potential of transcriptomics, the EPA issued a series of guidance, white papers, and reports between 2004 – 2007 that identified areas of application and barriers to implementation. The key barriers included a technical framework for genomic data analysis, criteria for acceptance of transcriptomics data, and consistency in methods for interpreting and analyzing transcriptomics data. Since that time, the technology and analysis methods for characterizing transcriptomic responses have matured and many of the barriers have been addressed through additional research and initiatives such as the MAQC/SEQC studies on reproducibility (MAQC 2006; SEQC 2014), the NTP consensus report on transcriptomic dose response modeling (NTP 2018), and OECD reporting template (Harrill et al. 2021b).

To evaluate the state of the science for using transcriptomics in quantitative human health assessment, a literature review was conducted comparing the concordance between transcriptomic PODs from short-term *in vivo* studies in rodents with apical PODs from traditional *in vivo* toxicity studies. The literature survey identified over 140 chemicals with diverse properties tested in 33 independent studies with varying experimental designs. The results of the literature survey demonstrated that transcriptomic BMD and BMDL values, when integrated at a gene set level, were consistently concordant with BMD and BMDL values for apical responses in traditional subchronic and chronic rodent toxicity studies. The transcriptomic and apical dose concordance was robust across different exposure durations, exposure routes, species, sex, target tissues, physicochemical properties, toxicokinetic half-lives, and technology platforms. For the 40 chemicals with reported chronic rodent bioassays results, the Pearson's correlation coefficient was 0.820 with a $\log_{10}$ RMSD of 0.593 ($\log_{10}$ mg/kg-day) and a median absolute ratio of 2.4 ± 1.0 (MAD). The RMSD value is similar to the range of inter-study standard deviation estimates for the LOAELs for systemic toxicity in repeated dose studies, approximated as residual RMSE in log10-mg/kg-day units [0.45-0.56; (Pham et al. 2020)]. The results suggest that the error associated with the concordance between the transcriptomic BMD values versus non-cancer and cancer apical BMD values is approximately equivalent to the inter-study variability in the repeated dose toxicity study itself.

In order to apply transcriptomics to human health assessment, a defined study design and data analysis process was evaluated to derive transcriptomic PODs based on recommendations in

the peer-reviewed National Toxicology Program Approach to Genomic Dose Response Modeling report (NTP 2018). The study design utilizes a 5-day, repeated dose *in vivo* study in male and female rats with an extended dose response series. The transcriptomic dose response modeling follows a stepwise process that utilizes BMD modeling approaches that are commonly employed in chemical risk assessment. A transcriptomic dose response modeling process was identified that resulted in a Pearson correlation coefficient and $\log_{10}$ RMSD for the transcriptomic and chronic apical BMD values of 0.910 and 0.567, respectively. The median absolute ratio of the transcriptomic BMD and chronic apical BMD values was 3.2 $\pm$ 1.9 (MAD). The inter-study transcriptomic BMD standard deviation for a subset of chemicals that were independently replicated was 0.242 ($\log_{10}$ mg/kg-day). The estimated family-wise error rate for identifying a gene set level BMD was 0.006. A statistical analysis demonstrated that the error associated with the concordance between the transcriptomic BMD values versus non-cancer and cancer apical BMD values was approximately equivalent to the combined inter-study variability associated with the 5-day transcriptomic study and the two-year rodent bioassay. The performance of the method in approximating an apical POD from two-year toxicity studies, the inter-study variability, and low family-wise error rate support the application of the defined study design and data analysis process to quantitative human health assessment.

There are caveats and limitations in the application of transcriptomic dose response to human health assessment and the analysis undertaken in this report. First, the use of gene sets for estimating the POD does not provide a direct measurement or link to a specific adversity. Rather, the POD must be interpreted as the dose at which there were no coordinated transcriptional changes that would indicate a potential toxicity of concern. The coordinated transcriptional changes used to identify the POD also do not necessarily discriminate between specific hazards, nor are they used to infer a mechanism or mode of action. Additional time points and analysis would be required to build a weight of evidence for a mode of action. Second, the coordinated transcriptional changes associated with the POD may be an adaptive response to chemical treatment rather than biological changes leading to an adverse endpoint. In the study by Gwinn and colleagues, two botanicals, ginseng and milk thistle extract, and tetrabromobisphenol A were not labeled as having either non-cancer or cancer-related lesions in male rats in the chronic bioassay, but the treatments did induce transcriptional changes and transcriptomic BMD values in the liver and kidney (Gwinn et al. 2020). However, in the original NTP reports, tetrabromobisphenol A was reported to have a reduction in body weight at the middle and high dose levels (NTP 2014) and milk thistle extract showed a reduction in biliary hyperplasia and mixed cell infiltration at the high dose (NTP 2011). While the histopathological changes following treatment with milk thistle extract may not be considered adverse in a traditional assessment context, apical changes were present that may have contributed to the transcriptional responses. Third, the use of the 5-day, repeat dose study with adult animals may not be predictive or protective of other health domains, such as developmental toxicity, reproductive toxicity, immunotoxicity, and neurotoxicity. Although this is also true of the traditional studies designed to assess systemic toxicity, a previous retrospective analysis demonstrated that

apical data from subchronic toxicity studies can predict effects on fertility (Dent 2007), while a separate retrospective analysis showed less than a two-fold difference in NOAEL values between rat subchronic studies and two-generation reproductive and developmental studies (Janer et al. 2007). Given the concordance between transcriptomic PODs from short-term studies and apical PODs from subchronic and chronic toxicity studies, these two retrospective analyses suggest that transcriptomic PODs may also be protective of apical PODs for other toxic effects, but further studies would be necessary to more confidently establish the degree of protection. Finally, the 14 chemicals and associated inter-study replicates used to refine the transcriptomic dose response methods may not be a random, representative sample of the chemicals that may be evaluated by ETAP in the future. Given the selection and relatively limited number of chemicals, the results may provide an overly optimistic assessment of the performance of the method when applied to new chemicals. However, the concordance associated with the 14 chemicals was similar to that identified for the 40 chemicals from the literature survey with chronic rodent BMD results, which suggests that the findings are a reasonable estimate. In addition, the number of chemicals used to refine the transcriptomic dose response methods and assess performance is similar to that used in the validation of multiple guideline assays and enhancements (Gelbke et al. 2007; Owens and Koeter 2003; Owens et al. 2007).

The overall conclusions from the literature survey, evaluation of the transcriptomic dose response analysis methods, and the statistical comparison of the concordance with inter-study variances support the use of transcriptomic PODs from 5-day, repeated dose *in vivo* rodent studies in quantitative human health assessments. The historical barriers that thus far limited application of transcriptomics in regulatory decision-making have mostly been addressed and the methods have undergone extensive peer-review in the individual publications and NTP report. While certain caveats and limitations remain, EPA is proposing to apply these methods in a standardized human health assessment framework to address the substantial data gaps that exist among chemicals that lack traditional toxicity testing data (EPA 2024b).

6. APPENDIX

6.1. SOURCE AND CALCULATION OF THE CHRONIC APICAL BMD AND BMDL VALUES

Apical BMD and BMDL values were obtained from the original Gwinn et al. publication (Gwinn et al. 2020) for twelve of the fourteen chemicals analyzed. For the remaining two chemicals, tris(2-chloroisopropyl) phosphate and di(2-ethylhexyl) phthalate, the histopathological results, feed, and body weight data for the male rat were extracted from the supplemental tables posted for peer review[28] and the technical report (NTP 2021). The BMD modeling was performed on the dichotomous cancer and non-cancer histopathological data using Benchmark Dose Modeling Software (BMDS) version 3.201. The doses were adjusted for study specific feed consumption and body weights. The incidence rates were poly 3 adjusted using values from the supplemental tables. A standard benchmark response (BMR) of 10% extra risk was used for all datasets. The following BMD models were fit to the data: Dichotomous Hill, Gamma, Log-Logistic, Weibull, Logistic, Log-Probit, Probit, Multi-stage 1°, Multi-stage 2°, Multi-stage 3°, and Multi-stage 4°. Per the default recommendations in the BMDS software, the Dichotomous Hill, Gamma, Log-Logistic, Multi-stage, and Weibull models were restricted. Selection of BMDs were based on the recommendation procedures described in the EPA BMD Technical Guidance (EPA 2012).

6.1.1. TRIS(2-CHLOROISOPROPYL) PHOSPHATE

The daily doses from the feeding study were calculated based on the average food consumption and body weights across all 102 weeks of the study. The average daily dose was 164.4, 339.7, 724.9, and 1433.8 mg/kg-day for the 2500, 5000, 10000, and 20000 ppm feed concentrations, respectively. For non-neoplastic endpoints, lung granulomatous focal inflammation was selected as the non-neoplastic endpoint due to the fact that there was a significant increase at the 725 and 1433.8 mg/kg-day doses (poly 3 test), and it had the lowest BMD. The incidence and poly-3 adjusted animal numbers are provided in Table 6-1.

Table 6-1. Daily dose, incidence, and poly 3 adjusted rates for lung granulomatous focal inflammation in the male rat exposed to tris(2-chloroisopropyl) phosphate for 102 weeks.

Daily Dose (mg/kg-day)	Number of Animals Based on Poly 3 Adjustment for Intercurrent Mortality	Incidence
0	39.93	7

[28] The chronic two-year bioassay results for BMD modeling were obtained from the NTP histopathological tables posted for peer review: https://ntp.niehs.nih.gov/go/TS-m20263

Table 6-1. Daily dose, incidence, and poly 3 adjusted rates for lung granulomatous focal inflammation in the male rat exposed to tris(2-chloroisopropyl) phosphate for 102 weeks.		
164	43.68	14
340	46.13	11
725	44.51	19
1434	47.08	24

The recommended model by BMDS would not be preferred for this analysis (Table 6-2). Rather, the log-logistic model was chosen due to the lower AIC since the purpose was to get the best estimate of the BMD (*i.e.*, central tendency) for comparing the concordance of the transcriptional and apical BMD values and not necessarily the "lowest BMDL" for a chemical assessment. The BMD, BMDL, and BMDU for lung granulomatous focal inflammation were 239.7, 140.7, and 834.4 mg/kg-day, respectively.

Table 6-2: BMD modeling output for lung granulomatous focal inflammation in the male rat.

Model	**BMD (mg/kg-day)**	**BMDL (mg/kg-day)**	**BMDU (mg/kg-day)**	**P Value**	**AIC**	**BMDS Recommendation**
Dichotomous Hill	223.6	58.8	743.2	0.374	296.07	Viable - Recommended
Gamma	303.1	203.9	768.5	0.535	294.27	Viable - Alternate
Log-Logistic	251.1	154.3	747.5	0.570	294.10	Viable - Alternate
Multistage Degree 4	303.1	203.9	798.8	0.535	294.27	Viable - Alternate
Multistage Degree 3	303.1	203.9	787.9	0.535	294.27	Viable - Alternate
Multistage Degree 2	303.1	203.9	768.1	0.535	294.27	Viable - Alternate
Multistage Degree 1	303.0	203.9	546.5	0.535	294.27	Viable - Alternate
Weibull	303.1	203.9	771.6	0.535	294.27	Viable - Alternate
Logistic	451.0	350.6	684.6	0.414	294.99	Viable - Alternate
Log-Probit	174.9	9.3	742.7	0.380	296.08	Questionable
Probit	433.9	335.5	668.4	0.428	294.90	Viable - Alternate

For the neoplastic lesions, malignant and benign tumors for all organs showed a significant increase at the 725 mg/kg-day dose (poly 3 test). However, the malignant and benign tumors (all organs) endpoint was not selected for modeling given the high rate in control animals. The lung granulomatous focal inflammation was chosen as the critical effect.

6.1.2. DI(2-ETHYLHEXYL) PHTHALATE

The NTP technical report identified pancreatic acinar adenoma or carcinoma (combined) in the postweaning-only study (Study 2) as the neoplasm in male rats with the lowest BMD. The average daily doses from the feeding study, incidence, and number of animals based on poly 3 adjustment for intercurrent mortality were listed in the NTP technical report. The doses, incidence, and poly 3 adjusted animal numbers are listed in Table 6-3.

Table 6-3. Daily dose, incidence, and poly 3 adjusted rates pancreatic acinar adenoma or carcinoma (combined) in the male rat exposed to di(2-ethylhexyl) phthalate.

Daily Dose (mg/kg-day)	Number of Animals Based on Poly 3 Adjustment for Intercurrent Mortality	Tumor Incidence
0	41.41	1
17	44.68	5
54	46.66	5
170	46.1	23
602	47.28	33

The log-logistic model was identified as the recommended model as it had the lowest AIC and a fit p-value > 0.1. The BMD, BMDL, and BMDU were 31.2, 20.3, and 63.1 mg/kg-day, respectively. Although the EPA typically uses linear multistage models for tumor dose response analysis, the fit p-values for the multi-stage models were all < 0.1 resulting in questionable fits to the data. If a multi-stage model was used, the first-degree model would be the best candidate since the b2 for the second-degree model was bounded. The BMD, BMDL, and BMDU for the first-degree linear multi-stage model were 44.7, 35.0, and 61.1 mg/kg-day, respectively. Given that the log-logistic model resulted in lower BMD and BMDL values, had a lower AIC, and a fit p-value > 0.1, it was selected as the preferred model. The details of the models are provided in Table 6-4.

Table 6-4: BMD modeling output for pancreatic acinar adenoma or carcinoma (combined) in the male rat.

Model	BMD (mg/kg-day)	BMDL (mg/kg-day)	BMDU (mg/kg-day)	P Value	AIC	BMDS Recommendation
Dichotomous Hill	71.0	20.7	150.4	0.110	205.06	Viable - Alternate
Gamma	44.7	35.0	68.1	0.069	205.29	Questionable
Log-Logistic	31.2	20.3	63.1	0.131	204.63	Viable - Recommended
Multistage Degree 4	44.7	35.0	61.1	0.069	205.29	Questionable

Table 6-4: BMD modeling output for pancreatic acinar adenoma or carcinoma (combined) in the male rat.						
Multistage Degree 3	44.7	35.0	61.1	0.069	205.29	Questionable
Multistage Degree 2	44.7	35.0	61.1	0.069	205.29	Questionable
Multistage Degree 1	44.7	35.0	59.1	0.069	205.29	Questionable
Weibull	44.7	35.0	66.0	0.069	205.29	Questionable
Logistic	122.5	100.6	149.2	0.000	218.56	Questionable
Log-Probit	32.8	15.5	70.1	0.112	204.93	Viable - Alternate
Probit	115.8	96.9	139.4	0.000	217.55	Questionable

A BMD analysis of non-neoplastic lesions in the male rat for Study 1 and 2 was performed, but no BMDs (*i.e.*, central tendency; not the corresponding BMDL) were identified that were lower than the pancreatic acinar adenoma or carcinoma (combined) when the log-logistic model was utilized. The pancreatic acinar adenoma or carcinoma (combined) was chosen as the critical effect.

6.2. SEQUENCE ALIGNMENT, NORMALIZATION AND QUALITY CONTROL METHODS FOR RE-ANALYSIS OF THE NTP TRANSCRIPTOMIC STUDIES

The two NTP publications possessed transcriptomic measurements on a total of 1450 individual liver and kidney samples from 23 distinct studies conducted at NTP (Gwinn et al. 2020)(https://doi.org/10.22427/NTP-DATA-002-00099-0001-000-1). The transcriptomic measurements were performed using the BioSpyder TempO-Seq rat S1500+ platform. The raw sequencing reads (FASTQ files) were aligned to known probe sequences to compute a matrix of read counts for each probe in each sample, as described previously (Harrill et al. 2021a). Initial quality checks were performed post-alignment to identify samples with insufficient sequencing depth or input RNA to yield reliable results. Each FASTQ file was aligned to the TempO-Seq probe manifest using HISAT2 (Kim et al. 2015; Kim et al. 2019). The alignment results were then imported directly into SAMtools (Li et al. 2009) to compute probe-level counts for each individual FASTQ file. Samples were examined for the following quality statistics described previously (Harrill et al. 2021a) and those not meeting minimum quality standards were removed from the analysis:

- Samples with total uniquely aligned reads fewer than 10% of the median read depth were removed.
- Samples with < 50% of sequenced reads uniquely aligned to known probes were removed.
- $NCov_5$ was computed as the total # of probes with at least 5 reads in each sample. This quality statistic is known to vary by cell type/source tissue (Harrill et al. 2021a). Liver samples with $NCov_5$ < 1,400 and Kidney samples with $NCov_5$ < 1,600 were removed.

- N_{80} was computed as the minimum number of probes that capture 80% of total mapped reads in the sample. No specific threshold for this value was set, but this quality statistic was considered when removing outliers, see below.

Of the 1450 total samples, 6 individual samples (0.4 %) were removed from further analysis based on the QC criteria above (Table 6-5). All removed samples had < 50% of sequenced reads uniquely aligned to known probe sequences.

Table 6-5. Samples removed based on QC metrics.

Tissue	Animal ID	Chemical[a]	Dose Group	QC Issue
Kidney	105R	DE71	0.38 mg/kg-day	< 50% reads aligned
Kidney	114R	DE71	1.5 mg/kg-day	< 50% reads aligned
Kidney	116R	DE71	1.5 mg/kg-day	< 50% reads aligned
Kidney	118R	DE71	3 mg/kg-day	< 50% reads aligned
Kidney	127	DE71	50 mg/kg-day	< 50% reads aligned
Kidney	2823	BDCA	20 mg/kg-day	< 50% reads aligned

[a]DE-71, Pentabromodiphenyl ether; BDCA, Bromodichloroacetic acid.

Prior to performing downstream gene expression analysis across samples, TempO-seq probe counts for each sample were normalized to adjust for differences in sequencing depth. For each set of samples to be used for dose response modeling (*i.e.*, each tissue for each test chemical), raw probe counts for all samples (including matched controls) were normalized within each sample as follows:

- All probes with a mean read count < 5 were removed, as these probes/genes lack sufficient signal for reliable analysis.
- Each remaining probe was normalized to Counts Per Million (CPM) = probe count * 1,000,000 / sum of all remaining probe counts in sample
- CPM values were transformed to $\log_2$ scale with added pseudo-count of 1 to prevent taking log of zero counts and ensuring a positive value for dose response modeling.

To identify potential outlier samples, a principal component analysis (PCA) was performed on subsets of samples corresponding to either: 1) all samples corresponding to same chemical and tissue, including matched vehicle controls ("treatment PCA"); and 2) all available vehicle controls corresponding to the same tissue ("vehicle PCA"). Samples not meeting the sequencing quality metrics described above (*e.g.*, < 50% of uniquely aligned reads) were excluded prior to PCA analysis. Outlier samples were identified based on the following considerations:

- Individual samples separated from all remaining samples on either principal component #1 (PC1) or principal component #2 (PC2) by >2x the span of all other samples on the corresponding PC were considered strong outliers and removed from further analysis.
- Individual samples separated by <2x the range of all other samples were considered moderate outliers, and additional exclusion criteria were considered:

- Vehicle samples that appear as moderate outliers on both a treatment PCA and vehicle PCA were excluded unless multiple controls from the same group appeared as outliers.
- Moderate outlier samples with lower quality than corresponding tissue samples by one or more sequencing quality metrics (*e.g.*, percentage of uniquely mapped reads) were excluded.
- Samples that appear as moderate outliers in both PC1 and PC2 compared with other remaining samples were excluded.
- Moderate outlier samples, where distance from all corresponding replicates or similar dose groups is greater than the pairwise distances between all other replicates of the same or similar doses, were excluded.

When multiple outlier samples were present on the same PCA, they were only removed if each outlier sample corresponded to a different dose group, as these are unlikely to represent any reproducible dose-dependent effect. A total of 21 samples (1 %) were removed as outliers (Table 6-6).

Table 6-6. Outlier samples removed based on PCA of samples grouped by chemical and tissue.

Tissue	Animal ID	Chemical[a]	Dose Group
Kidney	304R	TCAB	Vehicle
Kidney	810	TCPP	37.5 mg/kg-day
Kidney	818	TCPP	150 mg/kg-day
Kidney	822	TCPP	300 mg/kg-day
Liver	1007	PFOA	0.156 mg/kg-day
Liver	1012	PFOA	0.3125 mg/kg-day
Liver	1015	PFOA	0.625 mg/kg-day
Liver	1020	PFOA	1.25 mg/kg-day
Liver	1021	PFOA	2.5 mg/kg-day
Liver	1311	Methyl eugenol	9.25 mg/kg-day
Kidney	1318	Methyl eugenol	37 mg/kg-day
Kidney	1419	Coumarin	25 mg/kg-day
Liver	1432	Coumarin	200 mg/kg-day
Liver	1630	BDCA	80 mg/kg-day
Kidney	1632	BDCA	80 mg/kg-day
Liver	1716	α,β-Thujone	6.25 mg/kg-day
Liver	1801	Furan	Vehicle
Liver	1808	Furan	0.125 mg/kg-day
Liver	1824	Furan	2 mg/kg-day
Kidney	2226	Furan	4 mg/kg-day
Liver	2816	BDCA	5 mg/kg-day

Table 6-6. Outlier samples removed based on PCA of samples grouped by chemical and tissue.
[a]TCAB, 3,3',4,4'-Tetrachloroazobenzene; TCPP, Tris(2-chloroisopropyl) phosphate; PFOA, Perfluorooctanoic acid; BDCA, Bromodichloroacetic acid.

6.3. SEARCH STRATEGY FOR LITERATURE REVIEW

The following process was used to identify and screen articles for inclusion in the transcriptomic dose response sections of the literature review.

1. An initial list of expertly curated papers was assembled that focused on transcriptomic dose response analysis with application of benchmark dose methods (Table 6-7).
2. The articles were imported into Abstract Sifter[29] (Version 7)(Baker et al. 2017).
3. After importing, the 'Like articles' option was selected to identify similar articles[30]. The initial search also included articles that used other molecular measures besides gene expression (e.g., microRNA) and exposures that used nanoparticles; however, these articles were subsequently excluded due to the focus of the ETAP on mRNA and chemicals, respectively.

Table 6-7. Initial articles used as the base query for the literature review.

PMID	Year	Title	Log	Log time	Number of Like Articles
18499655	2008	Genomic signatures and dose-dependent transitions in nasal epithelial responses to inhaled formaldehyde in the rat.	Articles like 18499655	8/19/2022 16:11	145
20801182	2010	Toxicogenomics and cancer risk assessment: a framework for key event analysis and dose response assessment for nongenotoxic carcinogens.	Articles like 20801182	8/19/2022 16:10	200
20884683	2010	Formaldehyde: integrating dosimetry, cytotoxicity, and genomics to understand dose-dependent transitions for an endogenous compound.	Articles like 20884683	8/19/2022 15:05	157
21097997	2011	Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment.	Articles like 21097997	8/19/2022 15:04	144

[29] The latest version of Abstract Sifter can be downloaded from the EPA Computational Toxicology data download site at: https://www.epa.gov/chemical-research/downloadable-computational-toxicology-data
[30] For explanation of how the 'Like articles' are selected through Pubmed see: https://pubmed.ncbi.nlm.nih.gov/help/#computation-of-similar-articles

Table 6-7. Initial articles used as the base query for the literature review.

21795629	2011	Concentration- and time-dependent genomic changes in the mouse urinary bladder following exposure to arsenate in drinking water for up to 12 weeks.	Articles like 21795629	8/19/2022 15:01	86
22305970	2012	Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: a five chemical case study.	Articles like 22305970	8/19/2022 15:01	73
23125180	2013	Cross-species transcriptomic analysis of mouse and rat lung exposed to chloroprene.	Articles like 23125180	8/19/2022 15:00	90
23596260	2013	Temporal concordance between apical and transcriptional points of departure for chemical risk assessment.	Articles like 23596260	8/19/2022 14:59	115
23970803	2013	Concordance of transcriptional and apical benchmark dose levels for conazole-induced liver effects in mice.	Articles like 23970803	8/19/2022 14:59	114
24183702	2014	Case study on the utility of hepatic global gene expression profiling in the risk assessment of the carcinogen furan.	Articles like 24183702	8/19/2022 14:52	163
24449422	2014	Dose response modeling of early molecular and cellular key events in the CAR-mediated hepatocarcinogenesis pathway.	Articles like 24449422	8/19/2022 14:51	109
24952340	2014	Time- and concentration-dependent genomic responses of the rat airway to inhaled nickel subsulfide.	Articles like 24952340	8/19/2022 14:50	102
24976557	2014	Transcriptional responses in the rat nasal epithelium following subchronic inhalation of naphthalene vapor.	Articles like 24976557	8/19/2022 14:49	137
25605026	2015	Comparison of toxicogenomics and traditional approaches to inform mode of action and points of departure in human health risk assessment of benzo[a]pyrene in drinking water.	Articles like 25605026	8/19/2022 14:49	186
26194646	2016	Toxicogenomic assessment of liver responses following subchronic exposure to furan in Fischer F344 rats.	Articles like 26194646	8/19/2022 14:48	133
26519955	2016	Dose and Effect Thresholds for Early Key Events in a PPARα-Mediated Mode of Action.	Articles like 26519955	8/19/2022 14:47	94
26979667	2016	Nano-risk Science: application of toxicogenomics in an adverse outcome pathway framework for risk	Articles like 26979667	8/19/2022 8:55	200

Table 6-7. Initial articles used as the base query for the literature review.

		assessment of multi-walled carbon nanotubes.			
27638505	2017	Hepatic transcriptomic alterations for N,N-dimethyl-p-toluidine (DMPT) and p-toluidine after 5-day exposure in rats.	Articles like 27638505	8/19/2022 8:55	114
27858113	2017	A framework for the use of single-chemical transcriptomics data in predicting the hazards associated with complex mixtures of polycyclic aromatic hydrocarbons.	Articles like 27858113	8/19/2022 8:54	132
27859739	2016	Transcriptomic responses in the oral cavity of F344 rats and B6C3F1 mice following exposure to Cr(VI): Implications for risk assessment.	Articles like 27859739	8/19/2022 8:53	129
27928627	2017	Recommended approaches in the application of toxicogenomics to derive points of departure for chemical risk assessment.	Articles like 27928627	8/19/2022 8:53	138
28123101	2017	Editor's Highlight: Application of Gene Set Enrichment Analysis for Identification of Chemically Induced, Biologically Relevant Transcriptomic Networks and Potential Utilization in Human Health Risk Assessment.	Articles like 28123101	8/19/2022 8:52	93
28403741	2017	Impact of Acrylamide on Calcium Signaling and Cytoskeletal Filaments in Testes from F344 Rat.	Articles like 28403741	8/19/2022 8:50	76
28606764	2017	Transcriptional profiling of male F344 rats suggests the involvement of calcium signaling in the mode of action of acrylamide-induced thyroid cancer.	Articles like 28606764	8/19/2022 8:49	77
28717101	2017	Mechanism-based risk assessment strategy for drug-induced cholestasis using the transcriptional benchmark dose derived by toxicogenomics.	Articles like 28717101	8/19/2022 8:48	75
28927277	2017	Benchmark Dose Modeling Estimates of the Concentrations of Inorganic Arsenic That Induce Changes to the Neonatal Transcriptome, Proteome, and Epigenome in a Pregnancy Cohort.	Articles like 28927277	8/19/2022 8:47	98
28973375	2017	Editor's Highlight: Comparative Dose response Analysis of Liver and Kidney Transcriptomic Effects of Trichloroethylene and Tetrachloroethylene in B6C3F1 Mouse.	Articles like 28973375	8/19/2022 8:47	175

Table 6-7. Initial articles used as the base query for the literature review.

30589522	2018	NTP Research Report on In Vivo Repeat Dose Biological Potency Study of Triphenyl Phosphate (CAS No. 115-86-6) in Male Sprague Dawley Rats (Hsd: Sprague Dawley SD) (Gavage Studies): Research Report 8	Articles like 30589522	8/19/2022 8:46	103
30594549	2019	Hepatic transcriptional dose response analysis of male and female Fischer rats exposed to hexabromocyclododecane.	Articles like 30594549	8/19/2022 8:46	118
32268158	2020	Short-term toxicogenomics as an alternative approach to chronic in vivo studies for derivation of points of departure: A case study in the rat with a triazole fungicide.	Articles like 32268158	8/19/2022 8:44	117
32384157	2020	A Rat Liver Transcriptomic Point of Departure Predicts a Prospective Liver or Non-liver Apical Point of Departure.	Articles like 32384157	8/19/2022 8:43	106
32492150	2020	Evaluation of 5-day In Vivo Rat Liver and Kidney With High-throughput Transcriptomics for Estimating Benchmark Doses of Apical Outcomes.	Articles like 32492150	8/19/2022 8:39	112
32642447	2020	Early microRNA indicators of PPARα pathway activation in the liver.	Articles like 32642447	8/19/2022 8:38	144
32679240	2020	Comparative toxicity and liver transcriptomics of legacy and emerging brominated flame retardants following 5-day exposure in the rat.	Articles like 32679240	8/19/2022 8:37	121
32904430	2020	Transcriptomic data from the rat liver after five days of exposure to legacy or emerging brominated flame retardants.	Articles like 32904430	8/19/2022 8:35	97
33217531	2021	A rat subchronic study transcriptional point of departure estimates a carcinogenicity study apical point of departure.	Articles like 33217531	8/19/2022 8:33	107
35194992	2022	Integration of Toxicogenomics and Physiologically Based Pharmacokinetic Modeling in Human Health Risk Assessment of Perfluorooctane Sulfonate.	Articles like 35194992	8/19/2022 8:32	109
35537365	2022	Harmonization of transcriptomic and methylomic analysis in environmental epidemiology studies for potential application in chemical risk assessment.	Articles like 35537365	8/19/2022 8:30	103

Table 6-7. Initial articles used as the base query for the literature review.

35596682	2022	A microRNA or messenger RNA point of departure estimates an apical endpoint point of departure in a rat developmental toxicity model.	Articles like 35596682	8/19/2022 8:29	93
24194394	2014	Comparison of microarrays and RNA-seq for gene expression analyses of dose response experiments	Articles like 24194394	8/19/2022 14:52	200
26313361	2015	Impact of Genomics Platform and Statistical Filtering on Transcriptional Benchmark Doses (BMD) and Multiple Approaches for Selection of Chemical Point of Departure (PoD)	Articles like 26313361	8/19/2022 14:48	171

4. All 'Like articles' were imported into Abstract Sifter and duplicates removed to result in 874 articles.
5. The TermMap function in Abstract Sifter was used to sort the papers from high to low. The terms used are listed in Table 6-8.

Table 6-8. TermMap functions and results from Pubmed 'like articles' query.

Map to this:	When you see this:	Article Count	Boost by Term
Benchmark dose	benchmark dose	170	3
Benchmark dose	dose response	33	3
Benchmark dose	BMD	147	1
Transcript	genomic	256	2
Transcript	transcript	382	1
Transcript	gene	744	1
Transcript	transcriptome	208	2
Risk	risk	310	1
Risk	risk assessment	256	2
Benchmark dose total:	Total:	195	
Transcript total:	Total:	753	
Risk total:	Total:	310	

6. Titles and abstracts were scanned for relevance by two reviewers. If any conflicts were identified, consensus was reached through discussion. Only *in vivo* studies using mice or rats were retained as likely relevant. A total of 81 papers were identified (Table 6-9).

Table 6-9. Likely relevant articles identified following initial screening.

	TermMap Terms			

Table 6-9. Likely relevant articles identified following initial screening.

PMID	Benchmark Dose	Risk	Transcript	Score	Year	Title
17449896	13	14	5	32	2007	A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure.
20884683	9	12	10	31	2010	Formaldehyde: integrating dosimetry, cytotoxicity, and genomics to understand dose-dependent transitions for an endogenous compound.
21097997	12	15	12	39	2011	Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment.
22305970	11	16	18	45	2012	Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: a five chemical case study.
23596260	15	16	23	54	2013	Temporal concordance between apical and transcriptional points of departure for chemical risk assessment.
23970803	18	13	11	42	2013	Concordance of transcriptional and apical benchmark dose levels for conazole-induced liver effects in mice.
20849870	11	18	10	39	2013	Use of genomic data in risk assessment case study: I. Evaluation of the dibutyl phthalate male reproductive development toxicity data set.
23146762	11	18	10	39	2013	Gene expression profiling to identify potentially relevant disease outcomes and support human health risk assessment for carbon black nanoparticle exposure.
24194394	15	12	19	46	2014	Comparison of microarrays and RNA-seq for gene expression analyses of dose response experiments.
24183702	13	16	11	40	2014	Case study on the utility of hepatic global gene expression profiling in the risk assessment of the carcinogen furan.

Table 6-9. Likely relevant articles identified following initial screening.

24976557	11	11	12	34	2014	Transcriptional responses in the rat nasal epithelium following subchronic inhalation of naphthalene vapor.
24952340	11	11	10	32	2014	Time- and concentration-dependent genomic responses of the rat airway to inhaled nickel subsulfide.
24449422	10	5	14	29	2014	Dose response modeling of early molecular and cellular key events in the CAR-mediated hepatocarcinogenesis pathway.
26671443	12	14	20	46	2016	BMDExpress Data Viewer - a visualization tool to analyze BMDExpress datasets.
27601323	12	15	13	40	2016	Transcriptional benchmark dose modeling: Exploring how advances in chemical risk assessment may be applied to the radiation field.
27859739	9	13	18	40	2016	Transcriptomic responses in the oral cavity of F344 rats and B6C3F1 mice following exposure to Cr(VI): Implications for risk assessment.
26377693	9	11	19	39	2016	Comparative transcriptomic analyses to scrutinize the assumption that genotoxic PAHs exert effects via a common mode of action.
26194646	12	0	21	33	2016	Toxicogenomic assessment of liver responses following subchronic exposure to furan in Fischer F344 rats.
26496743	10	12	10	32	2016	Transcriptional Profiling of Dibenzo[def,p]chrysene-induced Spleen Atrophy Provides Mechanistic Insights into its Immunotoxicity in MutaMouse.
27562560	14	0	18	32	2016	Editor's Highlight: Dose response Analysis of RNA-Seq Profiles in Archival Formalin-Fixed Paraffin-Embedded Samples.

Table 6-9. Likely relevant articles identified following initial screening.

28123101	11	15	22	48	2017	Editor's Highlight: Application of Gene Set Enrichment Analysis for Identification of Chemically Induced, Biologically Relevant Transcriptomic Networks and Potential Utilization in Human Health Risk Assessment.
27638505	12	12	20	44	2017	Hepatic transcriptomic alterations for N,N-dimethyl-p-toluidine (DMPT) and p-toluidine after 5-day exposure in rats.
28717101	12	14	11	37	2017	Mechanism-based risk assessment strategy for drug-induced cholestasis using the transcriptional benchmark dose derived by toxicogenomics.
27858113	11	11	14	36	2017	A framework for the use of single-chemical transcriptomics data in predicting the hazards associated with complex mixtures of polycyclic aromatic hydrocarbons.
28973375	0	12	21	33	2017	Editor's Highlight: Comparative Dose response Analysis of Liver and Kidney Transcriptomic Effects of Trichloroethylene and Tetrachloroethylene in B6C3F1 Mouse.
28927277	12	0	20	32	2017	Benchmark Dose Modeling Estimates of the Concentrations of Inorganic Arsenic That Induce Changes to the Neonatal Transcriptome, Proteome, and Epigenome in a Pregnancy Cohort.
29475067	9	11	15	35	2018	Transcriptional profiling of male CD-1 mouse lungs and Harderian glands supports the involvement of calcium signaling in acrylamide-induced tumors.
30321009	11	11	10	32	2018	NTP Research Report on National Toxicology Program Approach to Genomic Dose response Modeling: Research Report 5

Table 6-9. Likely relevant articles identified following initial screening.

29329100	13	15	0	28	2018	A Web-Based System for Bayesian Benchmark Dose Estimation.
30594549	12	12	22	46	2019	Hepatic transcriptional dose response analysis of male and female Fischer rats exposed to hexabromocyclododecane.
30329029	2	11	21	34	2019	BMDExpress 2: enhanced transcriptomic dose response analysis workflow.
32413060	15	13	17	45	2020	The sensitivity of transcriptomics BMD modeling to the methods used for microarray data normalization.
32268158	9	12	11	32	2020	Short-term toxicogenomics as an alternative approach to chronic in vivo studies for derivation of points of departure: A case study in the rat with a triazole fungicide.
31950985	12	0	19	31	2020	BMDx: a graphical Shiny application to perform Benchmark Dose analysis for transcriptomics data.
32687419	13	13	18	44	2021	Meta-analysis of transcriptomic datasets using benchmark dose modeling shows value in supporting radiation risk assessment.
32761065	13	0	17	30	2021	FastBMD: an online tool for rapid benchmark dose response analysis of transcriptomics data.
35537365	11	15	23	49	2022	Harmonization of transcriptomic and methylomic analysis in environmental epidemiology studies for potential application in chemical risk assessment.
35939396	13	12	11	36	2022	Evaluating the Influences of Confounding Variables on Benchmark Dose using a Case Study in the Field of Ionizing Radiation.
35151117	13	12	11	36	2022	A computational system for Bayesian benchmark dose estimation of genomic data in BBMD.

Table 6-9. Likely relevant articles identified following initial screening.

35194992	11	17	5	33	2022	Integration of Toxicogenomics and Physiologically Based Pharmacokinetic Modeling in Human Health Risk Assessment of Perfluorooctane Sulfonate.
23125180	11	0	17	28	2013	Cross-species transcriptomic analysis of mouse and rat lung exposed to chloroprene.
28606764	9	11	8	28	2017	Transcriptional profiling of male F344 rats suggests the involvement of calcium signaling in the mode of action of acrylamide-induced thyroid cancer.
33362981	13	12	0	25	2020	bmd: an R package for benchmark dose estimation.
18441342	9	13	5	27	2008	Automated quantitative dose response modeling and point of departure determination for large toxicogenomic and high-throughput screening data sets.
21795629	11	11	5	27	2011	Concentration- and time-dependent genomic changes in the mouse urinary bladder following exposure to arsenate in drinking water for up to 12 weeks.
32642447	16	0	9	25	2020	Early microRNA indicators of PPARα pathway activation in the liver.
23831126	9	11	5	25	2013	Gene batteries and synexpression groups applied in a multivariate statistical approach to dose response analysis of toxicogenomic data.
17682005	0	12	11	23	2007	Exposure to arsenic at levels found in U.S. drinking water modifies expression in the mouse lung.
17961223	13	0	11	24	2007	BMDExpress: a software tool for the benchmark dose analyses of genomic data.
35939275	13	0	10	23	2022	Benchmark dose modeling of transcriptional data: a systematic approach to identify best practices for study designs used in radiation research.

Table 6-9. Likely relevant articles identified following initial screening.

18499655	13	0	10	23	2008	Genomic signatures and dose-dependent transitions in nasal epithelial responses to inhaled formaldehyde in the rat.
25554681	0	11	12	23	2015	MWCNTs of different physicochemical properties cause similar inflammatory responses, but differences in transcriptional and histological markers of fibrosis in mouse lungs.
28495587	13	0	10	23	2017	Dose response analysis of epigenetic, metabolic, and apical endpoints after short-term exposure to experimental hepatotoxicants.
30589522	12	0	10	22	2018	NTP Research Report on In Vivo Repeat Dose Biological Potency Study of Triphenyl Phosphate (CAS No. 115-86-6) in Male Sprague Dawley Rats (Hsd: Sprague Dawley SD) (Gavage Studies): Research Report 8
26519955	12	0	10	22	2016	Dose and Effect Thresholds for Early Key Events in a PPARα-Mediated Mode of Action.
31202941	0	13	8	21	2019	Hexabromocyclododecane (HBCD): A case study applying tiered testing for human health risk assessment.
32904430	10	0	12	22	2020	Transcriptomic data from the rat liver after five days of exposure to legacy or emerging brominated flame retardants.
18549499	0	11	10	21	2008	Sources of variation in baseline gene expression levels from toxicogenomics study control animals across multiple laboratories.
20957103	0	5	16	21	2010	Practical application of toxicogenomics for profiling toxicant-induced biological perturbations.
29945496	0	5	16	21	2018	Pathogenesis of Renal Injury and Gene Expression Changes in the Male CD-1 Mouse Associated with Exposure to Empagliflozin.

Table 6-9. Likely relevant articles identified following initial screening.

32679240	0	0	20	20	2020	Comparative toxicity and liver transcriptomics of legacy and emerging brominated flame retardants following 5-day exposure in the rat.
25958198	4	0	15	19	2015	RNA-Seq versus oligonucleotide array assessment of dose-dependent TCDD-elicited hepatic gene expression in mice.
27928627	11	16	23	50	2017	Recommended approaches in the application of toxicogenomics to derive points of departure for chemical risk assessment.
17118914	0	0	16	16	2006	Microarray-based compendium of hepatic gene expression profiles for prototypical ADME gene-inducing compounds in rats and mice in vivo.
32258091	12	0	5	17	2020	Aristolochic Acid-Induced Genotoxicity and Toxicogenomic Changes in Rodents.
35083394	12	0	5	17	2022	ALOHA: Aggregated local extrema splines for high-throughput dose response analysis.
35176998	0	0	16	16	2022	Evaluating the cytotoxicity and pathogenicity of multi-walled carbon nanotube through weighted gene co-expression network analysis: a nanotoxicogenomics study.
15056800	0	0	14	14	2004	Bromobenzene-induced hepatotoxicity at the transcriptome level.
26979667	12	15	13	40	2016	Nano-risk Science: application of toxicogenomics in an adverse outcome pathway framework for risk assessment of multi-walled carbon nanotubes.
28403741	11	0	2	13	2017	Impact of Acrylamide on Calcium Signaling and Cytoskeletal Filaments in Testes from F344 Rat.

Table 6-9. Likely relevant articles identified following initial screening.

32492150	13	5	19	37	2020	Evaluation of 5-day In Vivo Rat Liver and Kidney with High-throughput Transcriptomics for Estimating Benchmark Doses of Apical Outcomes.
33217531	12	13	17	42	2021	A rat subchronic study transcriptional point of departure estimates a carcinogenicity study apical point of departure.
25605026	9	16	11	36	2015	Comparison of toxicogenomics and traditional approaches to inform mode of action and points of departure in human health risk assessment of benzo[a]pyrene in drinking water.
32384157	11	13	20	44	2020	A Rat Liver Transcriptomic Point of Departure Predicts a Prospective Liver or Non-liver Apical Point of Departure.
26313361	12	13	11	36	2015	Impact of Genomics Platform and Statistical Filtering on Transcriptional Benchmark Doses (BMD) and Multiple Approaches for Selection of Chemical Point of Departure (PoD).
31881176	11	14	11	36	2020	A toxicogenomic approach for the risk assessment of the food contaminant acetamide.
20801182	10	14	5	29	2010	Toxicogenomics and cancer risk assessment: a framework for key event analysis and dose response assessment for nongenotoxic carcinogens.
15834898	0	0	7	7	2005	Evaluation of the gene expression changes induced by 17-alpha-ethynyl estradiol in the immature uterus/ovaries of the rat using high density oligonucleotide arrays.
21624382	0	5	2	7	2011	Multi-walled carbon nanotube-induced gene expression in the mouse lung: association with lung pathology.
35596682	2	11	11	24	2022	A microRNA or messenger RNA point of departure estimates an apical endpoint point of departure in a rat developmental toxicity model.

Table 6-9. Likely relevant articles identified following initial screening.

32449777	2	0	16	18	2020	TinderMIX: Time-dose integrated modelling of toxicogenomics data.
36518475	17	5	11	33	2022	Case study: Targeted RNA-sequencing of aged formalin-fixed paraffin-embedded samples for understanding chemical mode of action.

7. The 81 papers identified as potentially relevant (see Table 6-9) were reviewed in greater depth. The inclusion criteria were as follows: mouse or rat, at least 3 dose levels, BMD modeling of gene expression data obtained after 1 to 90 days of exposure, and comparison to apical effects Some of the articles identified as likely relevant were methods papers and some were re-analysis of data generated from different, earlier studies. Of the 81 likely relevant articles reviewed, there was data from 29 independent studies.
8. A separate search of NTP reports was performed to identify additional studies. The search identified 4 additional studies/reports for the literature review (NIEHS 2022a, b, c, d).
9. For the concordance analysis and scatter plots (Figures 3-1 to 3-3), only studies reporting a transcriptomic BMD (obtained after 1 to 90 days exposure) and two-year, chronic apical BMD. For those papers that reanalyzed data, the approach most similar to those proposed in the ETAP methods (e.g., the gene set with the lowest median BMD value) was used resulting in 21 articles contributing to the concordance analysis and scatter plots. The specific articles used for each figure are provided in Tables 6-10, 6-11, and 6-12 and cited in the figure legends.

Table 6-10. Rat and mouse transcriptomic BMD(L) values for lowest gene set from 1 to 90-day exposures and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-1.

Chemical (CASRN)	Species and strain/ stock	Sex	Route	Gene Expression Timepoint (d)	Gene Expression Tissue	Lowest Transcriptomic BMD (BMDL)	Endpoint	Lowest Traditional Endpoint BMD (BMDL)	BMD(L) Units	Study PubMed ID
Paraformaldehyde (30525-89-4)	F344/NCrl rat	M	I	1	Nasal epithelium	1.6 (1.1)	Nasal epithelium cell proliferation	4.9 (3.8)	ppm	17449896
Paraformaldehyde (30525-89-4)	F344/ CrlBR rat	M	I	7	Nasal epithelium	3.7 (2.7)	Nasal epithelium cell proliferation	4.9 (3.8)	ppm	20884683
Paraformaldehyde (30525-89-4)	F344/ CrlBR rat	M	I	28	Nasal epithelium	4.0 (2.7)	Nasal epithelium cell proliferation	4.9 (3.8)	ppm	20884683
Paraformaldehyde (30525-89-4)	F344/ CrlBR rat	M	I	91	Nasal epithelium	2.7 (1.9)	Nasal epithelium cell proliferation	4.9 (3.8)	ppm	20884683
Fenofibrate (49562-28-9)	Rat	F	G	2	Liver	4.7 (3.3)	Hepatocellular carcinoma	57 (40)	mg/kg	20801182
Methapyrilene (135-23-9)	Hsd:SD	M	G	7	Liver	28.8 (11.2)	Hepatocellular carcinoma	7.0 (4.0)	mg/kg	20801182
1,4-Dichlorobenzene (106-46-7)	B6C3F1 mouse	F	G	91	Liver	65 (38)	Liver adenoma and carcinoma	218 (158)	mg/kg	21097997
Methylene chloride (75-09-2)	B6C3F1 mouse	F	I	91	Lung	1137 (694)	Lung adenoma and carcinoma	791 (632)	mg/m^3	21097997
Methylene chloride (75-09-2)	B6C3F1 mouse	F	I	91	Liver	1265 (728)	Liver adenoma and carcinoma	3544 (1931)	mg/m^3	21097997
Naphthalene (91-20-3)	B6C3F1 mouse	F	I	91	Lung	17 (10)	Lung adenoma and carcinoma	120 (92)	mg/m^3	21097997
Propylene glycol mono-t-butyl ether (57018-52-7)	B6C3F1 mouse	F	I	91	Liver	677 (388)	Liver adenoma and carcinoma	1774 (866)	mg/m^3	21097997
1,2,3-trichloropropane (96-18-4)	B6C3F1 mouse	F	G	91	Liver	7.4 (4.0)	Liver adenoma and carcinoma	2.8 (1.3)	mg/kg	21097997
β-Chloroprene (126-99-8)	B6C3F1/Crl mouse	F	I	5	Lung	1.8 (1.1)	Time-to-lung tumor	1.2 (0.88)	ppm	23125180
β-Chloroprene (126-99-8)	B6C3F1/ Crl mouse	F	I	15	Lung	5.9 (3.9)	Time-to-lung tumor	1.2 (0.88)	ppm	23125180
β-Chloroprene (126-99-8)	F344/NCrl rat	F	I	5	Lung	38 (24)	Lung tumor	102 (76)	ppm	23125180
β-Chloroprene (126-99-8)	F344/NCrl rat	F	I	15	Lung	13 (8.0)	Lung tumor	102 (76)	ppm	23125180
4,4′-Methylenebis (N,N-dimethyl) benzenamine (101-61-1)	F344 rat	F	F	5	Thyroid	17 (13)	Thyroid follicular cell carcinoma and adenoma	30 (22)	adjusted to mg/kg	23596260
4,4′-Methylenebis (N,N-dimethyl) benzenamine (101-61-1)	F344 rat	F	F	14	Thyroid	23 (14)	Thyroid follicular cell carcinoma and adenoma	30 (22)	adjusted to mg/kg	23596260
4,4′-Methylenebis (N,N-dimethyl) benzenamine (101-61-1)	F344 rat	F	F	28	Thyroid	20 (12)	Thyroid follicular cell carcinoma and adenoma	30 (22)	adjusted to mg/kg	23596260
4,4′-Methylenebis (N,N-dimethyl) benzenamine (101-61-1)	F344 rat	F	F	91	Thyroid	21 (12)	Thyroid follicular cell carcinoma and adenoma	30 (22)	adjusted to mg/kg	23596260

Table 6-10. Rat and mouse transcriptomic BMD(L) values for lowest gene set from 1 to 90-day exposures and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-1.

Hydrazobenzene (122-66-7)	F344 rat	M	F	5	Liver	7.0 (4.2)	Hepatocellular carcinomas and neoplastic liver nodules	3.0 (1.7)	adjusted to mg/kg	23596260
Hydrazobenzene (122-66-7)	F344 rat	M	F	14	Liver	7.8 (6.1)	Hepatocellular carcinomas and neoplastic liver nodules	3.0 (1.7)	adjusted to mg/kg	23596260
Hydrazobenzene (122-66-7)	F344 rat	M	F	28	Liver	6.3 (4.0)	Hepatocellular carcinomas and neoplastic liver nodules	3.0 (1.7)	adjusted to mg/kg	23596260
Hydrazobenzene (122-66-7)	F344 rat	M	F	91	Liver	3.6 (2.0)	Hepatocellular carcinomas and neoplastic liver nodules	3.0 (1.7)	adjusted to mg/kg	23596260
N-Nitrosodiphenylamine (86-30-6)	F344 rat	F	F	5	Bladder	173 (130)	Transitional cell carcinoma of the bladder	184 (138)	adjusted to mg/kg	23596260
N-Nitrosodiphenylamine (86-30-6)	F344 rat	F	F	14	Bladder	104 (66)	Transitional cell carcinoma of the bladder	184 (138)	adjusted to mg/kg	23596260
N-Nitrosodiphenylamine (86-30-6)	F344 rat	F	F	28	Bladder	162 (124)	Transitional cell carcinoma of the bladder	184 (138)	adjusted to mg/kg	23596260
N-Nitrosodiphenylamine (86-30-6)	F344 rat	F	F	91	Bladder	158 (122)	Transitional cell carcinoma of the bladder	184 (138)	adjusted to mg/kg	23596260
Cyproconazole (94361-06-5)	CD-1 mouse	M	F	30	Liver	NA (9.0)	Liver adenoma or carcinoma	NA (2.1)	adjusted to mg/kg	23970803
Epoxiconazole (133855-98-8)	CD-1 mouse	M	F	30	Liver	NA (16)	Liver adenoma or carcinoma	NA (33)	adjusted to mg/kg	23970803
Propiconazole (60207-90-1)	CD-1 mouse	M	F	30	Liver	NA (56)	Liver adenoma or carcinoma	NA (65)	adjusted to mg/kg	23970803
Triadimefon (43121-43-3)	CD-1 mouse	M	F	30	Liver	NA (58)	Liver adenoma or carcinoma	NA (270)	adjusted to mg/kg	23970803
Furan (110-00-9)	B3C6F1 mouse	F	G	21	Liver	1.9 (1.4)	Hepatocellular adenoma	2.3 (1.3)	mg/kg	24183702
Benzo(a)pyrene (50-32-8)	B6C3F1 mouse	M	G	3	Liver	0.3 (0.2)	Liver tumor	1.8 (1.2)	mg/kg	25605026
Benzo(a)pyrene (50-32-8)	Muta™ Mouse	M	G	28	Liver	25 (5.3)	Liver tumor	1.8 (1.2)	mg/kg	25605026
Furan (110-00-9)	F344 rat	M	G	90	Liver	0.08 (0.05)	Hepatocellular adenoma or carcinoma	1.8 (1.3)	mg/kg	26194646
Furan (110-00-9)	F344 rat	F	G	90	Liver	0.6 (0.4)	Hepatocellular adenoma or carcinoma	5.5 (3.9)	mg/kg	26194646
Di(2-ethylhexyl) phthalate (117-81-7)	B6C3F1 mouse	M	F	7	Liver	55 (38)	Hepatocellular adenoma or carcinoma	35 (NA)	adjusted to mg/kg	27562560
Coal tar extract (NA)	Muta™ Mouse	M	G	28	Lung	1.6 (1.0)	Lung tumor	2.5 (1.5)	mg/kg	27858113
Acrylamide (79-06-1)	F344/ DuCrl rat	M	DW	15	Thyroid	2.9 (0.68)	Thyroid tumors	1.4 (0.82)	mg/kg	28606764
Acrylamide (79-06-1)	F344/ DuCrl rat	M	DW	31	Thyroid	0.81 (0)	Thyroid tumors	1.4 (0.82)	mg/kg	28606764
Tetrachloroethylene (127-18-4)	B6C3F1 mouse	M	G	1	Kidney	NA (7.5)	Kidney adenoma and adenocarcinoma	NA (63)	HED mg/kg	28973375
Tetrachloroethylene (127-18-4)	B6C3F1 mouse	M	G	1	Liver	NA (139)	Liver adenoma and carcinoma	NA (48)	HED mg/kg	28973375

Table 6-10. Rat and mouse transcriptomic BMD(L) values for lowest gene set from 1 to 90-day exposures and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-1.

Trichloroethylene (79-01-6)	B6C3F1 mouse	M	G	1	Kidney	NA (0.029)	Increased kidney/body weight ratio	NA (0.69)	HED mg/kg	28973375
Trichloroethylene (79-01-6)	B6C3F1 mouse	M	G	1	Liver	NA (3.1)	Liver adenoma and carcinoma	NA (3.5)	HED mg/kg	28973375
Acrylamide (79-06-1)	F344/ DuCrl rat	M	DW	31	Testes	4.8 (3.3)	Peripheral neuropathy	NA (0.43)	mg/kg	28403741
Acrylamide (79-06-1)	CD-1 mouse	M	DW	15	Lung	2.9 (2.2)	Lung adenoma	2.1 (1.3)	mg/kg	29475067
Acrylamide (79-06-1)	CD-1 mouse	M	DW	31	Harderian gland	0.87 (0.54)	Harderian gland adenoma	0.37 (0.17)	mg/kg	29475067
Myclobutanil (88671-89-0)	Crl:CD (SD) rat	M	G	14	Testes	50 (25)	Testis seminiferous tubule atrophy (unilateral or bilateral)	2.9 (1.4)	mg/kg	32268158
Acrylamide (79-06-1)	Sprague-Dawley rat	M	G	5	Kidney	0.68 (NA)	Peripheral nerve (sciatic) axon degeneration-non neoplastic	0.61 (0.43)	mg/kg	32492150
Bromodichloroacetic acid (71133-14-7)	Sprague-Dawley rat	M	G	5	Kidney	1.3 (NA)	Bone marrow angiectasis-non neoplastic	2.3 (1.9)	mg/kg	32492150
Coumarin (91-64-5)	Sprague-Dawley rat	M	G	5	Liver	3.0 (NA)	Liver necrosis-non neoplastic	5.9 (4.9)	mg/kg	32492150
Pentabromodiphenyl ether mixture (32534-81-9)	Sprague-Dawley rat	M	G	5	Kidney	0.72 (NA)	Hepatocyte hypertrophy-non neoplastic	0.15 (0.11)	mg/kg	32492150
Di(2-ethylhexyl) phthalate (117-81-7)	Sprague-Dawley rat	M	G	5	Liver	23 (NA)	Testis degeneration seminiferous tubules-non neoplastic	184 (155)	mg/kg	32492150
Ethinyl estradiol (57-63-6)	Sprague-Dawley rat	M	G	5	Kidney	0.19 (NA)	Mammary gland alveolar hyperplasia-non neoplastic	0.69 (0.47)	µg/kg	32492150
Furan (110-00-9)	Sprague-Dawley rat	M	G	5	Kidney	0.61 (NA)	Liver cholangiofibrosis-non neoplastic	0.10 (0.09)	mg/kg	32492150
Hexachlorobenzene (118-74-1)	Sprague-Dawley rat	M	G	5	Liver	4.6 (NA)	Kidney chronic nephrosis (severe)-non neoplastic	0.59 (0.35)	mg/kg	32492150
Methyl eugenol (93-15-2)	Sprague-Dawley rat	M	G	5	Liver	33 (NA)	Hepatocellular adenoma or carcinoma	12 (9.9)	mg/kg	32492150
Perfluorooctanoic acid (335-67-1)	Sprague-Dawley rat	M	G	5	Kidney	0.11 (NA)	Hepatocyte hypertrophy-non neoplastic	0.50 (0.41)	mg/kg	32492150
Pulegone (89-82-7)	Sprague-Dawley rat	M	G	5	Liver	21 (NA)	Nasal olfactory epithelium degeneration-non neoplastic	14.2 (9.5)	mg/kg	32492150
3,3',4,4'-Tetrachloroazobenzene (14047-09-7)	Sprague-Dawley rat	M	G	5	Kidney	1.7 (NA)	Forestomach epithelium hyperplasia-non neoplastic	2.9 (2.13)	mg/kg	32492150
α,β-Thujone (76231-76-0)	Sprague-Dawley rat	M	G	5	Kidney	141 (NA)	Kidney mineralization-non neoplastic	3.2 (2.3)	mg/kg	32492150
Fenpicoxamid (517875-34-2)	Sprague-Dawley rat	M	F	90	Kidney	115 (59)	Kidney, chronic progressive glomerulonephropathy, slight	NA (140)	adjusted to mg/kg	33217531
Pronamide (23950-58-5)	F344/ DuCrl rat	M	F	90	Liver	5.6 (2.5)	Liver weight, relative	NA (8.5)	adjusted to mg/kg	33217531

Table 6-10. Rat and mouse transcriptomic BMD(L) values for lowest gene set from 1 to 90-day exposures and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-1.

Sulfoxaflor (946578-00-3)	F344/ DuCrl rat	M	F	90	Liver	6.1 (4.0)	Testis adenoma, unilateral, bilateral and unilateral combined	NA (1.0)	adjusted to mg/kg	33217531
Triclopyr Acid (55335-06-3)	F344/ DuCrl rat	M	F	90	Kidney	5.4 (3.0)	Relative kidney weight	NA (4.0)	adjusted to mg/kg	33217531
Dichloroacetic acid (79-43-6)	B6C3F1 mouse	M	DW	6	Liver	134 (101)	Hepatocellular adenoma or carcinoma	44 (13)	adjusted to mg/kg	36518475

Legend: **Sex**: (M) Male and (F) Female | **Exposure Route:** (DW) Drinking Water, (F) Feed, (G) Gavage, (I) Inhalation | **BMD:** Benchmark dose, **BMD(L):** Benchmark dose lower confidence limit, **CASRN**: Chemical Abstracts Service Registry Number, **d:** Day, **HED:** Human equivalent dose, **NA:** Not available

Table 6-11. Rat and mouse transcriptomic BMD(L) values for lowest gene set in liver and kidney tissue from 1 to 90-day exposures and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-2.

Chemical (CASRN)	Species and strain/ stock	Sex	Route	Gene Expression Timepoint (d)	Gene Expression Tissue	Lowest Transcriptomic BMD (BMDL)	Endpoint	Lowest Traditional Endpoint BMD (BMDL)	BMD(L) Units	Study PubMed ID
Acrylamide (79-06-1)	Sprague-Dawley rat	M	G	5	Liver	4.7 (NA)	Peripheral nerve (sciatic) axon degeneration-non neoplastic	0.61 (0.43)	mg/kg	32492150
Acrylamide (79-06-1)	Sprague-Dawley rat	M	G	5	Kidney	0.68 (NA)	Peripheral nerve (sciatic) axon degeneration-non neoplastic	0.61 (0.43)	mg/kg	32492150
Bromodichloroacetic acid (71133-14-7)	Sprague-Dawley rat	M	G	5	Liver	2.8 (NA)	Bone marrow angiectasis-non neoplastic	2.3 (1.9)	mg/kg	32492150
Bromodichloroacetic acid (71133-14-7)	Sprague-Dawley rat	M	G	5	Kidney	1.3 (NA)	Bone marrow angiectasis-non neoplastic	2.3 (1.9)	mg/kg	32492150
Coumarin (91-64-5)	Sprague-Dawley rat	M	G	5	Kidney	21 (NA)	Liver necrosis-non neoplastic	5.9 (4.9)	mg/kg	32492150
Pentabromodiphenyl ether mixture (32534-81-9)	Sprague-Dawley rat	M	G	5	Kidney	0.72 (NA)	Hepatocyte hypertrophy-non neoplastic	0.15 (0.11)	mg/kg	32492150
Di(2-ethylhexyl) phthalate (117-81-7)	Sprague-Dawley rat	M	G	5	Liver	23 (NA)	Testis degeneration seminiferous tubules-non neoplastic	184 (155)	mg/kg	32492150
Di(2-ethylhexyl) phthalate (117-81-7)	Sprague-Dawley rat	M	G	5	Kidney	44 (NA)	Testis degeneration seminiferous tubules-non neoplastic	184 (155)	mg/kg	32492150
Ethinyl estradiol (57-63-6)	Sprague-Dawley rat	M	G	5	Liver	1.8 (NA)	Mammary gland alveolar hyperplasia-non neoplastic	0.69 (0.47)	µg/kg	32492150
Ethinyl estradiol (57-63-6)	Sprague-Dawley rat	M	G	5	Kidney	0.19 (NA)	Mammary gland alveolar hyperplasia-non neoplastic	0.69 (0.47)	µg/kg	32492150
Furan (110-00-9)	Sprague-Dawley rat	M	G	5	Kidney	0.61 (NA)	Liver cholangiofibrosis-non neoplastic	0.10 (0.09)	mg/kg	32492150
Hexachlorobenzene (118-74-1)	Sprague-Dawley rat	M	G	5	Liver	4.6 (NA)	Kidney chronic nephrosis (severe)-non neoplastic	0.59 (0.35)	mg/kg	32492150
Methyl eugenol (93-15-2)	Sprague-Dawley rat	M	G	5	Kidney	116 (NA)	Hepatocellular adenoma or carcinoma	12 (9.9)	mg/kg	32492150
Perfluorooctanoic acid (335-67-1)	Sprague-Dawley rat	M	G	5	Kidney	0.11 (NA)	Hepatocyte hypertrophy-non neoplastic	0.50 (0.41)	mg/kg	32492150
Pulegone (89-82-7)	Sprague-Dawley rat	M	G	5	Liver	21 (NA)	Nasal olfactory epithelium degeneration-non neoplastic	14 (9.5)	mg/kg	32492150
Pulegone (89-82-7)	Sprague-Dawley rat	M	G	5	Kidney	22(NA)	Nasal olfactory epithelium degeneration-non neoplastic	14 (9.5)	mg/kg	32492150
3,3',4,4'-Tetrachloroazobenzene (14047-09-7)	Sprague-Dawley rat	M	G	5	Liver	12 (NA)	Forestomach epithelium hyperplasia-non neoplastic	2.9 (2.1)	mg/kg	32492150
3,3',4,4'-Tetrachloroazobenzene (14047-09-7)	Sprague-Dawley rat	M	G	5	Kidney	1.7 (NA)	Forestomach epithelium hyperplasia-non neoplastic	2.9 (2.1)	mg/kg	32492150

Table 6-11. Rat and mouse transcriptomic BMD(L) values for lowest gene set in liver and kidney tissue from 1 to 90-day exposures and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-2.

α,β-Thujone (76231-76-0)	Sprague-Dawley rat	M	G	5	Liver	151 (NA)	Kidney mineralization-non neoplastic	3.2 (2.3)	mg/kg	32492150
Tetrachloroethylene (127-18-4)	B6C3F1 mouse	M	G	1	Liver	NA (139)	Kidney adenoma and adenocarcinoma	NA (63)	HED mg/kg	28973375
Tetrachloroethylene (127-18-4)	B6C3F1 mouse	M	G	1	Kidney	NA (7.5)	Liver adenoma and carcinoma	NA (48)	HED mg/kg	28973375
Trichloroethylene (79-01-6)	B6C3F1 mouse	M	G	1	Liver	NA (3.1)	Increased kidney/body weight ratio	NA (0.69)	HED mg/kg	28973375
Trichloroethylene (79-01-6)	B6C3F1 mouse	M	G	1	Kidney	NA (0.029)	Liver adenoma and carcinoma	NA (3.5)	HED mg/kg	28973375
Methylene chloride (75-09-2)	B6C3F1 mouse	F	I	91	Liver	1265 (728)	Lung adenoma and carcinoma	791 (632)	mg/m^3	21097997
Acrylamide (79-06-1)	F344/DuCrl rat	M	DW	31	Liver	3.8 (1.9)	Thyroid tumors	1.4 (0.82)	mg/kg	28606764
Sulfoxaflor (946578-00-3)	F344/DuCrl rat	M	F	90	Liver	6.1 (4.0)	Testis adenoma, unilateral, bilateral and unilateral combined	NA (1.0)	adjusted to mg/kg	33217531
Triclopyr Acid (55335-06-3)	F344/DuCrl rat	M	F	90	Liver	31 (9.3)	Relative kidney weight	NA (4.0)	adjusted to mg/kg	33217531

Legend: **Sex**: (M) Male and (F) Female | **Exposure Route:** (DW) Drinking Water, (F) Feed, (G) Gavage, (I) Inhalation | **BMD:** Benchmark dose, **BMD(L):** Benchmark dose lower confidence limit, **CASRN**: Chemical Abstracts Service Registry Number, **d:** Day, **HED:** Human equivalent dose, **NA:** Not available

Table 6-12. Rat and mouse transcriptomic BMD(L) values for the lowest gene set from 1 to 90-day exposures collected using microarray, TempO-Seq, and RNA-seq platforms and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-3.

Chemical (CASRN)	Species and strain/ stock	Sex	Platform	Route	Gene Expression Timepoint (d)	Gene Expression Tissue	Lowest Transcriptomic BMD (BMDL)	Endpoint	Lowest Traditional Endpoint BMD (BMDL)	BMD(L) Units	Study PubMed ID
Paraformaldehyde (30525-89-4)	F344/ NCrl rat	M	Array	I	1	Nasal epithelium	1.6 (1.1)	Nasal epithelium cell proliferation	4.9 (3.8)	ppm	17449896
Paraformaldehyde (30525-89-4)	F344/ CrlBR rat	M	Array	I	7	Nasal epithelium	3.7 (2.7)	Nasal epithelium cell proliferation	4.9 (3.8)	ppm	20884683
Paraformaldehyde (30525-89-4)	F344/ CrlBR rat	M	Array	I	28	Nasal epithelium	4.0 (2.7)	Nasal epithelium cell proliferation	4.9 (3.8)	ppm	20884683
Paraformaldehyde (30525-89-4)	F344/ CrlBR rat	M	Array	I	91	Nasal epithelium	2.7 (1.9)	Nasal epithelium cell proliferation	4.9 (3.8)	ppm	20884683
Fenofibrate (49562-28-9)	Rat	F	Array	G	2	Liver	4.7 (3.3)	Hepatocellular carcinoma	57 (40)	mg/kg	20801182
Methapyrilene (135-23-9)	Sprague-Dawley rat	M	Array	G	7	Liver	29 (11)	Hepatocellular carcinoma	7.0 (4.0)	mg/kg	20801182
1,4-Dichlorobenzene (106-46-7)	B6C3F1 mouse	F	Array	G	91	Liver	65 (38)	Liver adenoma and carcinoma	218 (158)	mg/kg	21097997
Methylene chloride (75-09-2)	B6C3F1 mouse	F	Array	I	91	Lung	1137 (694)	Lung adenoma and carcinoma	791 (632)	mg/m3	21097997
Methylene chloride (75-09-2)	B6C3F1 mouse	F	Array	I	91	Liver	1265 (728)	Liver adenoma and carcinoma	3545 (1931)	mg/m3	21097997
Naphthalene (91-20-3)	B6C3F1 mouse	F	Array	I	91	Lung	17 (10)	Lung adenoma and carcinoma	120 (92)	mg/m3	21097997
Propylene glycol mono-t-butyl ether (57018-52-7)	B6C3F1 mouse	F	Array	I	91	Liver	677 (388)	Liver adenoma and carcinoma	1774 (866)	mg/m3	21097997
1,2,3-trichloropropane (96-18-4)	B6C3F1 mouse	F	Array	G	91	Liver	7.4 (4.0)	Liver adenoma and carcinoma	2.8 (1.3)	mg/kg	21097997
β-Chloroprene (126-99-8)	B6C3F1/ Crl mouse	F	Array	I	5	Lung	1.8 (1.1)	Time-to-lung tumor	1.2 (0.88)	ppm	23125180
β-Chloroprene (126-99-8)	B6C3F1/ Crl mouse	F	Array	I	15	Lung	5.9 (3.9)	Time-to-lung tumor	1.2 (0.88)	ppm	23125180
β-Chloroprene (126-99-8)	F344/ NCrl rat	F	Array	I	5	Lung	38 (24)	Lung tumor	102 (76)	ppm	23125180
β-Chloroprene (126-99-8)	F344/ NCrl rat	F	Array	I	15	Lung	12 (8.0)	Lung tumor	102 (76)	ppm	23125180
4,4′-Methylenebis (N,N-dimethyl) benzenamine (101-61-1)	F344 rat	F	Array	F	5	Thyroid	17 (13)	Thyroid follicular cell carcinoma and adenoma	30.1 (22.4)	adjusted to mg/kg	23596260

Table 6-12. Rat and mouse transcriptomic BMD(L) values for the lowest gene set from 1 to 90-day exposures collected using microarray, TempO-Seq, and RNA-seq platforms and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-3.

4,4′-Methylenebis (N,N-dimethyl) benzenamine (101-61-1)	F344 rat	F	Array	F	14	Thyroid	23 (14)	Thyroid follicular cell carcinoma and adenoma	30 (22)	adjusted to mg/kg	23596260
4,4′-Methylenebis (N,N-dimethyl) benzenamine (101-61-1)	F344 rat	F	Array	F	28	Thyroid	20 (11.9)	Thyroid follicular cell carcinoma and adenoma	30 (22)	adjusted to mg/kg	23596260
4,4′-Methylenebis (N,N-dimethyl) benzenamine (101-61-1)	F344 rat	F	Array	F	91	Thyroid	21 (12)	Thyroid follicular cell carcinoma and adenoma	30 (22)	adjusted to mg/kg	23596260
Hydrazobenzene (122-66-7)	F344 rat	M	Array	F	5	Liver	7.0 (4.2)	Hepatocellular carcinomas and neoplastic liver nodules	3 (1.7)	adjusted to mg/kg	23596260
Hydrazobenzene (122-66-7)	F344 rat	M	Array	F	14	Liver	7.8 (6.1)	Hepatocellular carcinomas and neoplastic liver nodules	3.0 (1.7)	adjusted to mg/kg	23596260
Hydrazobenzene (122-66-7)	F344 rat	M	Array	F	28	Liver	6.3 (4.0)	Hepatocellular carcinomas and neoplastic liver nodules	3.0 (1.7)	adjusted to mg/kg	23596260
Hydrazobenzene (122-66-7)	F344 rat	M	Array	F	91	Liver	3.6 (2.0)	Hepatocellular carcinomas and neoplastic liver nodules	3.0 (1.7)	adjusted to mg/kg	23596260
N-Nitrosodiphenylamine (86-30-6)	F344 rat	F	Array	F	5	Bladder	173 (130)	Transitional cell carcinoma of the bladder	184 (138)	adjusted to mg/kg	23596260
N-Nitrosodiphenylamine (86-30-6)	F344 rat	F	Array	F	14	Bladder	104 (66)	Transitional cell carcinoma of the bladder	184 (138)	adjusted to mg/kg	23596260
N-Nitrosodiphenylamine (86-30-6)	F344 rat	F	Array	F	28	Bladder	162 (124)	Transitional cell carcinoma of the bladder	184 (138)	adjusted to mg/kg	23596260
N-Nitrosodiphenylamine (86-30-6)	F344 rat	F	Array	F	91	Bladder	158 (122)	Transitional cell carcinoma of the bladder	184 (138)	adjusted to mg/kg	23596260
Cyproconazole (94361-06-5)	CD-1 mouse	M	Array	F	30	Liver	NA (9.0)	Liver adenoma or carcinoma	NA (2.1)	adjusted to mg/kg	23970803
Epoxiconazole (133855-98-8)	CD-1 mouse	M	Array	F	30	Liver	NA (16)	Liver adenoma or carcinoma	NA (33)	adjusted to mg/kg	23970803
Propiconazole (60207-90-1)	CD-1 mouse	M	Array	F	30	Liver	NA (56)	Liver adenoma or carcinoma	NA (65)	adjusted to mg/kg	23970803
Triadimefon (43121-43-3)	CD-1 mouse	M	Array	F	30	Liver	NA (58)	Liver adenoma or carcinoma	NA (270)	adjusted to mg/kg	23970803

Table 6-12. Rat and mouse transcriptomic BMD(L) values for the lowest gene set from 1 to 90-day exposures collected using microarray, TempO-Seq, and RNA-seq platforms and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-3.

Furan (110-00-9)*	B3C6F1 mouse	F	Array	G	21	Liver	1.9 (1.4)	Hepatocellular adenoma	2.3 (1.3)	mg/kg	24183702
Benzo(a)pyrene (50-32-8)	B6C3F1 mouse	M	Array	G	3	Liver	0.30 (0.20)	Liver tumor	1.8 (1.2)	mg/kg	25605026
Benzo(a)pyrene (50-32-8)	Muta™Mouse	M	Array	G	28	Liver	25 (5.3)	Liver tumor	1.8 (1.2)	mg/kg	25605026
Furan (110-00-9)*	B6C3F1 mouse	F	RNA-Seq	G	21	Liver	1.7 (1.0)	Hepatocellular adenoma or carcinoma	2.3 (1.3)	mg/kg	26313361
Furan (110-00-9)	F344 rat	M	Array	G	90	Liver	0.08 (0.05)	Hepatocellular adenoma or carcinoma	1.8 (1.3)	mg/kg	26194646
Furan (110-00-9)	F344 rat	F	Array	G	90	Liver	0.60 (0.40)	Hepatocellular adenoma or carcinoma	5.5 (3.9)	mg/kg	26194646
Di(2-ethylhexyl) phthalate (117-81-7)	B6C3F1 mouse	M	RNA-Seq	F	7	Liver	55 (38)	Hepatocellular adenoma or carcinoma	35 (NA)	adjusted to mg/kg	27562560
Coal tar extract (NA)	Muta™Mouse	M	Array	G	28	Lung	1.6 (1.0)	Lung tumor	2.5 (1.5)	mg/kg	27858113
Acrylamide (79-06-1)	F344/DuCrl rat	M	RNA-Seq	DW	15	Thyroid	2.9 (0.7)	Thyroid tumors	1.4 (0.82)	mg/kg	28606764
Acrylamide (79-06-1)	F344/DuCrl rat	M	RNA-Seq	DW	31	Thyroid	0.81 (0)	Thyroid tumors	1.4 (0.82)	mg/kg	28606764
Tetrachloroethylene (127-18-4)	B6C3F1 mouse	M	RNA-Seq	G	1	Kidney	NA (7.5)	Kidney adenoma and adenocarcinoma	NA (63)	HED mg/kg	28973375
Tetrachloroethylene (127-18-4)	B6C3F1 mouse	M	RNA-Seq	G	1	Liver	NA (139)	Liver adenoma and carcinoma	NA (48)	HED mg/kg	28973375
Trichloroethylene (79-01-6)	B6C3F1 mouse	M	RNA-Seq	G	1	Kidney	NA (0.029)	Increased kidney/body weight ratio	NA (0.69)	HED mg/kg	28973375
Trichloroethylene (79-01-6)	B6C3F1 mouse	M	RNA-Seq	G	1	Liver	NA (3.1)	Liver adenoma and carcinoma	NA (3.5)	HED mg/kg	28973375
Acrylamide (79-06-1)	F344/DuCrl rat	M	RNA-Seq	DW	31	Testes	4.8 (3.3)	Preputial gland: duct ectasia	0.11 (0.015)	mg/kg	28403741
Acrylamide (79-06-1)	CD-1 mouse	M	RNA-Seq	DW	15	Lung	2.9 (2.2)	Lung adenoma	2.1 (1.3)	mg/kg	29475067
Acrylamide (79-06-1)	CD-1 mouse	M	RNA-Seq	DW	31	Harderian gland	0.87 (0.54)	Harderian gland adenoma	0.37 (0.17)	mg/kg	29475067
Myclobutanil (88671-89-0)	Crl:CD (SD) rat	M	RNA-Seq	G	14	Testes	50 (25)	Testis seminiferous tubule atrophy (unilateral or bilateral)	2.9 (1.4)	mg/kg	32268158
Acrylamide (79-06-1)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	4.7 (NA)	Peripheral nerve (sciatic) axon degeneration-non neoplastic	0.61 (0.43)	mg/kg	32492150
Acrylamide (79-06-1)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	0.68 (NA)	Peripheral nerve (sciatic) axon degeneration-non neoplastic	0.61 (0.43)	mg/kg	32492150

Table 6-12. Rat and mouse transcriptomic BMD(L) values for the lowest gene set from 1 to 90-day exposures collected using microarray, TempO-Seq, and RNA-seq platforms and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-3.

Bromodichloroacetic acid (71133-14-7)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	2.8 (NA)	Bone marrow angiectasis-non neoplastic	2.3 (1.9)	mg/kg	32492150
Bromodichloroacetic acid (71133-14-7)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	1.3 (NA)	Bone marrow angiectasis-non neoplastic	2.3 (1.9)	mg/kg	32492150
Coumarin (91-64-5)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	3.0 (NA)	Liver necrosis-non neoplastic	5.9 (4.9)	mg/kg	32492150
Coumarin (91-64-5)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	21 (NA)	Liver necrosis-non neoplastic	5.9 (4.9)	mg/kg	32492150
Pentabromodiphenyl ether mixture (32534-81-9)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	10 (NA)	Hepatocyte hypertrophy-non neoplastic	0.15 (0.11)	mg/kg	32492150
Pentabromodiphenyl ether mixture (32534-81-9)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	0.72 (NA)	Hepatocyte hypertrophy-non neoplastic	0.15 (0.11)	mg/kg	32492150
Di(2-ethylhexyl) phthalate (117-81-7)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	23 (NA)	Testis degeneration seminiferous tubules-non neoplastic	184 (155)	mg/kg	32492150
Di(2-ethylhexyl) phthalate (117-81-7)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	44 (NA)	Testis degeneration seminiferous tubules-non neoplastic	184 (155)	mg/kg	32492150
Ethinyl estradiol (57-63-6)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	1.8 (NA)	Mammary gland alveolar hyperplasia-non neoplastic	0.69 (0.47)	ug/kg	32492150
Ethinyl estradiol (57-63-6)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	0.19 (NA)	Mammary gland alveolar hyperplasia-non neoplastic	0.69 (0.47)	µg/kg	32492150
Furan (110-00-9)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	1.2 (NA)	Liver cholangiofibrosis-non neoplastic	0.1 (0.09)	mg/kg	32492150
Furan (110-00-9)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	0.61 (NA)	Liver cholangiofibrosis-non neoplastic	0.10 (0.09)	mg/kg	32492150
Hexachlorobenzene (118-74-1)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	4.6 (NA)	Kidney chronic nephrosis (severe)-non neoplastic	0.59 (0.35)	mg/kg	32492150
Hexachlorobenzene (118-74-1)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	4.6 (NA)	Adrenal gland adrenal pheochromocytoma	1.2 (0.60)	mg/kg	32492150
Methyl eugenol (93-15-2)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	33 (NA)	Hepatocellular adenoma or carcinoma	12 (9.9)	mg/kg	32492150
Methyl eugenol (93-15-2)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	116 (NA)	Hepatocellular adenoma or carcinoma	12 (9.9)	mg/kg	32492150

Table 6-12. Rat and mouse transcriptomic BMD(L) values for the lowest gene set from 1 to 90-day exposures collected using microarray, TempO-Seq, and RNA-seq platforms and BMD(L) values for traditional non-cancer and cancer endpoints from rodent bioassays reported in literature search. The data used to create Figure 3-3.

Perfluorooctanoic acid (335-67-1)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	0.35 (NA)	Hepatocyte hypertrophy-non neoplastic	0.50 (0.41)	mg/kg	32492150
Perfluorooctanoic acid (335-67-1)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	0.11 (NA)	Hepatocyte hypertrophy-non neoplastic	0.50 (0.41)	mg/kg	32492150
Pulegone (89-82-7)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	21 (NA)	Nasal olfactory epithelium degeneration-non neoplastic	14 (9.5)	mg/kg	32492150
Pulegone (89-82-7)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	22 (NA)	Nasal olfactory epithelium degeneration-non neoplastic	14 (9.5)	mg/kg	32492150
3,3',4,4'-Tetrachloroazobenzene (14047-09-7)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	12 (NA)	Forestomach epithelium hyperplasia-non neoplastic	2.9 (2.1)	mg/kg	32492150
3,3',4,4'-Tetrachloroazobenzene (14047-09-7)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	1.7 (NA)	Forestomach epithelium hyperplasia-non neoplastic	2.9 (2.1)	mg/kg	32492150
α,β-Thujone (76231-76-0)	Sprague-Dawley rat	M	TempO-Seq	G	5	Liver	151 (NA)	Kidney mineralization-non neoplastic	3.2 (2.3)	mg/kg	32492150
α,β-Thujone (76231-76-0)	Sprague-Dawley rat	M	TempO-Seq	G	5	Kidney	141 (NA)	Kidney mineralization-non neoplastic	3.2 (2.3)	mg/kg	32492150
Fenpicoxamid (517875-34-2)	Sprague-Dawley rat	M	RNA-Seq	F	90	Kidney	115 (59)	Kidney, chronic progressive glomerulonephropathy, slight	NA (140)	adjusted to mg/kg	33217531
Pronamide (23950-58-5)	F344/Du Crl rat	M	RNA-Seq	F	90	Liver	5.6 (2.5)	Liver weight, relative	NA (8.5)	adjusted to mg/kg	33217531
Sulfoxaflor (946578-00-3)	F344/Du Crl rat	M	RNA-Seq	F	90	Liver	6.1 (4.0)	Testis adenoma, unilateral, bilateral and unilateral combined	NA (1.0)	adjusted to mg/kg	33217531
Triclopyr Acid (55335-06-3)	F344/Du Crl rat	M	RNA-Seq	F	90	Liver	31 (9.3)	Relative kidney weight	NA (4.0)	adjusted to mg/kg	33217531
Triclopyr Acid (55335-06-3)	F344/Du Crl rat	M	RNA-Seq	F	90	Kidney	5.4 (3.0)	Relative kidney weight	NA (4.0)	adjusted to mg/kg	33217531
Dichloroacetic acid (79-43-6)*	B6C3F1 mouse	M	RNA-Seq	DW	6	Liver	94 (74)	Hepatocellular adenoma or carcinoma	44 (13)	adjusted to mg/kg	36518475
Dichloroacetic acid (79-43-6)*	B6C3F1 mouse	M	TempO-Seq	DW	6	Liver	134 (101)	Hepatocellular adenoma or carcinoma	44 (13)	adjusted to mg/kg	36518475

Legend: **Sex**: (M) Male and (F) Female | **Exposure Route:** (DW) Drinking Water, (F) Feed, (G) Gavage, (I) Inhalation | **Platform:** (Array) Microarray | **BMD:** Benchmark dose, **BMD(L):** Benchmark dose lower confidence limit, **CASRN**: Chemical Abstracts Service Registry Number, **d:** Day, **HED:** Human equivalent dose, **NA:** Not available**;** *Paired data; the same samples have conventional RNA-Seq and TempO-Seq data.

7. REFERENCES

Andersen ME, Clewell HJ, 3rd, Bermudez E, Willson GA, Thomas RS. 2008. Genomic signatures and dose-dependent transitions in nasal epithelial responses to inhaled formaldehyde in the rat. Toxicol Sci 105:368-383.

Andersen ME, Clewell HJ, 3rd, Bermudez E, Dodd DE, Willson GA, Campbell JL, et al. 2010. Formaldehyde: integrating dosimetry, cytotoxicity, and genomics to understand dose-dependent transitions for an endogenous compound. Toxicol Sci 118:716-731.

Baker N, Knudsen T, Williams A. 2017. Abstract Sifter: a comprehensive front-end system to PubMed. F1000Res 6.

Barupal DK, Fiehn O. 2019. Generating the Blood Exposome Database Using a Comprehensive Text Mining and Database Fusion Approach. Environ Health Perspect 127:97008.

Bercu JP, Jolly RA, Flagella KM, Baker TK, Romero P, Stevens JL. 2010. Toxicogenomics and cancer risk assessment: a framework for key event analysis and dose-response assessment for nongenotoxic carcinogens. Regul Toxicol Pharmacol 58:369-381.

Bhat VS, Hester SD, Nesnow S, Eastmond DA. 2013. Concordance of transcriptional and apical benchmark dose levels for conazole-induced liver effects in mice. Toxicol Sci 136:205-215.

Bianchi E, Costa E, Yan ZJ, Murphy L, Howell J, Anderson D, et al. 2021. A rat subchronic study transcriptional point of departure estimates a carcinogenicity study apical point of departure. Food Chem Toxicol 147:111869.

Black MB, Parks BB, Pluta L, Chu TM, Allen BC, Wolfinger RD, et al. 2014. Comparison of microarrays and RNA-seq for gene expression analyses of dose-response experiments. Toxicol Sci 137:385-403.

Bushel PR, Paules RS, Auerbach SS. 2018. A Comparison of the TempO-Seq S1500+ Platform to RNA-Seq and Microarray Using Rat Liver Mode of Action Samples. Front Genet 9:485.

Cannizzo MD, Wood CE, Hester SD, Wehmas LC. 2022. Case study: Targeted RNA-sequencing of aged formalin-fixed paraffin-embedded samples for understanding chemical mode of action. Toxicol Rep 9:883-894.

Catlin NR, Collins BJ, Auerbach SS, Ferguson SS, Harnly JM, Gennings C, et al. 2018. How similar is similar enough? A sufficient similarity case study with Ginkgo biloba extract. Food Chem Toxicol 118:328-339.

Chepelev NL, Moffat ID, Labib S, Bourdon-Lacombe J, Kuo B, Buick JK, et al. 2015. Integrating toxicogenomics into human health risk assessment: lessons learned from the benzo[a]pyrene case study. Crit Rev Toxicol 45:44-52.

Chepelev NL, Gagné R, Maynor T, Kuo B, Hobbs CA, Recio L, et al. 2017. Transcriptional profiling of male F344 rats suggests the involvement of calcium signaling in the mode of action of acrylamide-induced thyroid cancer. Food Chem Toxicol 107:186-200.

Chepelev NL, Gagné R, Maynor T, Kuo B, Hobbs CA, Recio L, et al. 2018. Transcriptional profiling of male CD-1 mouse lungs and Harderian glands supports the involvement of calcium signaling in acrylamide-induced tumors. Regul Toxicol Pharmacol 95:75-90.

Clewell HJ, Thomas RS, Kenyon EM, Hughes MF, Adair BM, Gentry PR, et al. 2011. Concentration- and time-dependent genomic changes in the mouse urinary bladder following exposure to arsenate in drinking water for up to 12 weeks. Toxicol Sci 123:421-432.

Clewell HJ, Efremenko A, Campbell JL, Dodd DE, Thomas RS. 2014. Transcriptional responses in the rat nasal epithelium following subchronic inhalation of naphthalene vapor. Toxicol Appl Pharmacol 280:78-85.

Danforth C, Chiu WA, Rusyn I, Schultz K, Bolden A, Kwiatkowski C, et al. 2020. An integrative method for identification and prioritization of constituents of concern in produced water from onshore oil and gas extraction. Environ Int 134:105280.

Dent MP. 2007. Strengths and limitations of using repeat-dose toxicity studies to predict effects on fertility. Regul Toxicol Pharmacol 48:241-258.

Dong H, Gill S, Curran IH, Williams A, Kuo B, Wade MG, et al. 2016. Toxicogenomic assessment of liver responses following subchronic exposure to furan in Fischer F344 rats. Arch Toxicol 90:1351-1367.

Dunnick JK, Shockley KR, Morgan DL, Brix A, Travlos GS, Gerrish K, et al. 2017. Hepatic transcriptomic alterations for N,N-dimethyl-p-toluidine (DMPT) and p-toluidine after 5-day exposure in rats. Arch Toxicol 91:1685-1696.

EEA. 2018. Chemicals for a Sustainable Future – Report of the EEA Scientific Committee Seminar. 978-92-9213-936-0. Luxembourg:European Environment Agency.

EPA. 2004a. Acetochlor Report of the Cancer Assessment Review Committee (CARC). Washington, D.C.

EPA. 2004b. Potential Implications of Genomics for Regulatory and Risk Assessment Applications at EPA. EPA 100/B-04/002. Washington, DC.

EPA. 2007. Interim Guidance for Microarray-Based Assays: Data Submission, Quality, Analysis, Data Management, and Training Considerations External Review Draft. Washington, DC.

EPA. 2009. An Approach To Using Toxicogenomic Data in U.S. EPA Human Health Risk Assessments: A Dibutyl Phthalate Case Study. EPA/600/R-09/028F. Washington, DC:U.S. Environmental Protection Agency.

EPA. 2012. Benchmark Dose Technical Guidance. EPA/100/R-12/001. Washington, DC:U.S. Environmental Protection Agency.

EPA. 2021. New Approach Methods Work Plan. EPA/600/X-21/209. Washington, DC:U.S. Environmental Protection Agency.

EPA. 2022. ORD Staff Handook for Developing IRIS Assessments. EPA 600/R-22/268. Research Triangle Park, NC:U.S. Environmental Protection Agency.

EPA. 2024a. Value of Information Case Study on the Human Health and Economic Trade-offs Associated with the Timeliness, Uncertainty, and Costs of the Draft EPA Transcriptomic Assessment Product (ETAP). EPA/600/X-23/082. Research Triangle Park, NC:U.S. Environmental Protection Agency.

EPA. 2024b. Standard Methods for Development of EPA Transcriptomic Assessment Products (ETAPs). EPA/600/X-23/083. Research Triangle Park, NC:U.S. Environmental Protection Agency.

Farmahin R, Williams A, Kuo B, Chepelev NL, Thomas RS, Barton-Maclaren TS, et al. 2017. Recommended approaches in the application of toxicogenomics to derive points of departure for chemical risk assessment. Arch Toxicol 91:2045-2065.

Foox J, Tighe SW, Nicolet CM, Zook JM, Byrska-Bishop M, Clarke WE, et al. 2021. Performance assessment of DNA sequencing platforms in the ABRF Next-Generation Sequencing Study. Nat Biotechnol 39:1129-1140.

Fowler T, Sen R, Roy AL. 2011. Regulation of primary response genes. Mol Cell 44:348-360.

Gelbke HP, Hofmann A, Owens JW, Freyberger A. 2007. The enhancement of the subacute repeat dose toxicity test OECD TG 407 for the detection of endocrine active chemicals: comparison with toxicity tests of longer duration. Arch Toxicol 81:227-250.

Geter DR, Bhat VS, Gollapudi BB, Sura R, Hester SD. 2014. Dose-response modeling of early molecular and cellular key events in the CAR-mediated hepatocarcinogenesis pathway. Toxicol Sci 138:425-445.

Gwinn WM, Auerbach SS, Parham F, Stout MD, Waidyanatha S, Mutlu E, et al. 2020. Evaluation of 5-day In Vivo Rat Liver and Kidney With High-throughput Transcriptomics for Estimating Benchmark Doses of Apical Outcomes. Toxicol Sci 176:343-354.

Hagiwara S, Paoli GM, Price PS, Gwinn MR, Guiseppi-Elie A, Farrell PJ, et al. 2022. A value of information framework for assessing the trade-offs associated with uncertainty, duration, and cost of chemical toxicity testing. Risk Anal 43:498-515.

Harrill JA, Everett LJ, Haggard DE, Sheffield T, Bundy JL, Willis CM, et al. 2021a. High-Throughput Transcriptomics Platform for Screening Environmental Chemicals. Toxicol Sci 181:68-89.

Harrill JA, Viant MR, Yauk CL, Sachana M, Gant TW, Auerbach SS, et al. 2021b. Progress towards an OECD reporting framework for transcriptomics and metabolomics in regulatory toxicology. Regul Toxicol Pharmacol 125:105020.

Hester S, Eastmond DA, Bhat VS. 2015. Developing toxicogenomics as a research tool by applying benchmark dose-response modelling to inform chemical mode of action and tumorigenic potency. Int J of Biotechnol 14:28-46.

Hester SD, Bhat V, Chorley BN, Carswell G, Jones W, Wehmas LC, et al. 2016. Editor's Highlight: Dose-Response Analysis of RNA-Seq Profiles in Archival Formalin-Fixed Paraffin-Embedded Samples. Toxicol Sci 154:202-213.

Isaacs KK, Wall JT, Williams AR, Hobbie KA, Sobus JR, Ulrich E, et al. 2022. A harmonized chemical monitoring database for support of exposure assessments. Sci Data 9:314.

Jackson AF, Williams A, Recio L, Waters MD, Lambert IB, Yauk CL. 2014. Case study on the utility of hepatic global gene expression profiling in the risk assessment of the carcinogen furan. Toxicol Appl Pharmacol 274:63-77.

Janer G, Hakkert BC, Piersma AH, Vermeire T, Slob W. 2007. A retrospective analysis of the added value of the rat two-generation reproductive toxicity study versus the rat subchronic toxicity study. Reprod Toxicol 24:103-113.

Johnson KJ, Auerbach SS, Costa E. 2020. A Rat Liver Transcriptomic Point of Departure Predicts a Prospective Liver or Non-liver Apical Point of Departure. Toxicol Sci 176:86-102.

Johnson KJ, Auerbach SS, Stevens T, Barton-Maclaren TS, Costa E, Currie RA, et al. 2022. A Transformative Vision for an Omics-Based Regulatory Chemical Testing Paradigm. Toxicol Sci 190:127-132.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12:357-360.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907-915.

Krewski D, Andersen ME, Tyshenko MG, Krishnan K, Hartung T, Boekelheide K, et al. 2020. Toxicity testing in the 21st century: progress in the past decade and future perspectives. Arch Toxicol 94:1-58.

Labib S, Williams A, Guo CH, Leingartner K, Arlt VM, Schmeiser HH, et al. 2016. Comparative transcriptomic analyses to scrutinize the assumption that genotoxic PAHs exert effects via a common mode of action. Arch Toxicol 90:2461-2480.

Labib S, Williams A, Kuo B, Yauk CL, White PA, Halappanavar S. 2017. A framework for the use of single-chemical transcriptomics data in predicting the hazards associated with complex mixtures of polycyclic aromatic hydrocarbons. Arch Toxicol 91:2599-2616.

Lake AD, Wood CE, Bhat VS, Chorley BN, Carswell GK, Sey YM, et al. 2016. Dose and Effect Thresholds for Early Key Events in a PPARα-Mediated Mode of Action. Toxicol Sci 149:312-325.

LaRocca J, Costa E, Sriram S, Hannas BR, Johnson KJ. 2020. Short-term toxicogenomics as an alternative approach to chronic in vivo studies for derivation of points of departure: A case study in the rat with a triazole fungicide. Regul Toxicol Pharmacol 113:104655.

Li D, Kusko R, Ning B, Tong W, Johann DJ, Jr., Xu J. 2021. FDA-led consortium studies advance quality control of targeted next generation sequencing assays for precision oncology. Precis Cancer Med 4.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079.

Li S, Tighe SW, Nicolet CM, Grove D, Levy S, Farmerie W, et al. 2014. Multi-platform assessment of transcriptome profiling using RNA-seq in the ABRF next-generation sequencing study. Nat Biotechnol 32:915-925.

Lobenhofer EK, Cui X, Bennett L, Cable PL, Merrick BA, Churchill GA, et al. 2004. Exploration of low-dose estrogen effects: identification of No Observed Transcriptional Effect Level (NOTEL). Toxicol Pathol 32:482-492.

MAQC. 2006. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 24:1151-1161.

Mav D, Shah RR, Howard BE, Auerbach SS, Bushel PR, Collins JB, et al. 2018. A hybrid gene selection approach to create the S1500+ targeted gene sets for use in high-throughput transcriptomics. PLoS One 13:e0191105.

Mercer TR, Xu J, Mason CE, Tong W, on behalf of the MSC. 2021. The Sequencing Quality Control 2 study: establishing community standards for sequencing in precision medicine. Genome Biol 22:306.

Moffat I, Chepelev N, Labib S, Bourdon-Lacombe J, Kuo B, Buick JK, et al. 2015. Comparison of toxicogenomics and traditional approaches to inform mode of action and points of departure in human health risk assessment of benzo[a]pyrene in drinking water. Crit Rev Toxicol 45:1-43.

NASEM. 2007. Toxicity Testing in the 21st Century: A Vision and a Strategy. Washington, D.C.:National Academies of Sciences, Engineering, and Medicine.

NASEM. 2009. Science and Decisions: Advancing Risk Assessment. Washington, DC:National Academies of Sciences, Engineering, and Medicine.

National Toxicology P. 2018. NTP Research Reports. In: NTP Research Report on In Vivo Repeat Dose Biological Potency Study of Triphenyl Phosphate (CAS No 115-86-6) in Male Sprague Dawley Rats (Hsd: Sprague Dawley SD) (Gavage Studies): Research Report 8. Durham (NC):National Toxicology Program.

NIEHS. 2022a. NIEHS Report on the In Vivo Repeat Dose Biological Potency Study of Tricresyl Phosphate (CASRN 1330-78-5) in Male Sprague Dawley (Hsd:Sprague Dawley® SD®) Rats (Gavage Studies). NIEHS Report 06. Research Triangle Park, NC:U.S. Department of Health and Human Services.

NIEHS. 2022b. NIEHS Report on the In Vivo Repeat Dose Biological Potency Study of 2-Ethylhexyl Diphenyl Phosphate (CASRN 1241-94-7) in Male Sprague Dawley (Hsd:Sprague Dawley® SD®) Rats (Gavage Studies). NIEHS Report 02. Research Triangle Park, NC.

NIEHS. 2022c. NIEHS Report on In Vivo Repeat Dose Biological Potency Study of tert-Butylphenyl Diphenyl Phosphate (CASRN 56803-37-3) in Male Sprague Dawley (Hsd:Sprague Dawley® SD®) Rats(Gavage Studies) NIEHS Report 03 Research Triangle Park, NC.

NIEHS. 2022d. NIEHS Report on the In Vivo Repeat Dose Biological Potency Study of Isodecyl Diphenyl Phosphate (CASRN 29761-21-5) in Male Sprague Dawley (Hsd:Sprague Dawley® SD®) Rats (Gavage Studies). NIEHS Report 04. Research Triangle Park, NC.

NTP. 2011. Toxicology and Carcinogenesis Studies of Milk Thistle Extract (CASRN 84604-20-6) in F344/N Rats and B6C3F1 Mice (Feed studies). TR-565. Research Triangle Park, NC:National Toxicology Program.

NTP. 2014. Toxicology Studies of Tetrabromobisphenol A in F344/NTac Rats and B6C3F1/N Mice and Toxicology and Carcinogenesis Studies of Tetrabromobisphenol A in Wistar Han [Crl:WI(Han)] Rats and B6C3F1/N Mice (Gavage Studies). TR-587. Research Triangle Park, NC:National Toxicology Program.

NTP. 2018. NTP Research Report on National Toxicology Program Approach to Genomic Dose-Response Modeling. NTP-RR-5. Research Triangle Park, NC:National Toxicology Program, U.S. Department of Health and Human Services.

NTP. 2021. NTP Technical Report on the Toxicology and Carcinogenesis Studies of Di(2-ethylhexyl) Phthalate (CASRN 117-81-7) Administered in Feed to Sprague Dawley (HSD:Sprague Dawley SD) Rats. NTP TR 601. Research Triangle Park, NC:National Toxicology Program.

NTP. 2022. Chemical Effects in Biological Systems (CEBS). Research Triangle Park, USA:National Toxicology Program.

OE. 2019. The Global Chemical Industry: Catalyzing Growth and Addressing Our World's Sustainability Challenges. Washington, DC:Oxford Economics.

OECD. 2018. Toward a New Comprehensive Global Database of Per- and polyfluoroalkyl substances (PFASs): Summary Report on Updating the OeCD 2007 List of Per- and Polyfluoroalkyl Substances (PFASs). ENV/JM/MONO(2018)7.Organization for Economic Cooperation and Development.

Owens W, Koeter HB. 2003. The OECD program to validate the rat uterotrophic bioassay: an overview. Environ Health Perspect 111:1527-1529.

Owens W, Gray LE, Zeiger E, Walker M, Yamasaki K, Ashby J, et al. 2007. The OECD program to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses: phase 2 dose-response studies. Environ Health Perspect 115:671-678.

Pham LL, Watford S, Pradeep P, Martin MT, Thomas R, Judson R, et al. 2020. Variability in in vivo studies: Defining the upper limit of performance for predictions of systemic effect levels. Comput Toxicol 15:1-100126.

Phillips JR, Svoboda DL, Tandon A, Patel S, Sedykh A, Mav D, et al. 2019. BMDExpress 2: enhanced transcriptomic dose-response analysis workflow. Bioinformatics 35:1780-1782.

Recio L, Friedman M, Marroni D, Maynor T, Chepelev NL. 2017. Impact of Acrylamide on Calcium Signaling and Cytoskeletal Filaments in Testes From F344 Rat. Int J Toxicol 36:124-132.

Rowlands JC, Budinsky R, Gollapudi B, Black MB, Wolfinger RD, Cukovic D, et al. 2013. A genomics-based analysis of relative potencies of dioxin-like compounds in primary rat hepatocytes. Toxicol Sci 136:595-604.

SEQC. 2014. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. Nat Biotechnol 32:903-914.

SEQC/MAQC-III. 2014. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. Nat Biotechnol 32:903-914.

Shockley KR, Cora MC, Malarkey DE, Jackson-Humbles D, Vallant M, Collins BJ, et al. 2020. Comparative toxicity and liver transcriptomics of legacy and emerging brominated flame retardants following 5-day exposure in the rat. Toxicol Lett 332:222-234.

Thomas RS, Allen BC, Nong A, Yang L, Bermudez E, Clewell HJ, 3rd, et al. 2007. A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. Toxicol Sci 98:240-248.

Thomas RS, Clewell HJ, 3rd, Allen BC, Wesselkamper SC, Wang NC, Lambert JC, et al. 2011. Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. Toxicol Sci 120:194-205.

Thomas RS, Clewell HJ, 3rd, Allen BC, Yang L, Healy E, Andersen ME. 2012. Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: a five chemical case study. Mutat Res 746:135-143.

Thomas RS, Himmelstein MW, Clewell HJ, 3rd, Yang Y, Healy E, Black MB, et al. 2013a. Cross-species transcriptomic analysis of mouse and rat lung exposed to chloroprene. Toxicol Sci 131:629-640.

Thomas RS, Wesselkamper SC, Wang NC, Zhao QJ, Petersen DD, Lambert JC, et al. 2013b. Temporal concordance between apical and transcriptional points of departure for chemical risk assessment. Toxicol Sci 134:180-194.

Tullai JW, Schaffer ME, Mullenbrock S, Sholder G, Kasif S, Cooper GM. 2007. Immediate-early and delayed primary response genes are distinct in function and genomic architecture. J Biol Chem 282:23981-23995.

Wang Z, Walker GW, Muir DCG, Nagatani-Yoshida K. 2020. Toward a Global Understanding of Chemical Pollution: A First Comprehensive Analysis of National and Regional Chemical Inventories. Environ Sci Technol 54:2575-2584.

Yang L, Allen BC, Thomas RS. 2007. BMDExpress: a software tool for the benchmark dose analyses of genomic data. BMC Genomics 8:387.

Yeakley JM, Shepard PJ, Goyena DE, VanSteenhouse HC, McComb JD, Seligmann BE. 2017. A trichostatin A expression signature identified by TempO-Seq targeted whole transcriptome profiling. PLoS One 12:e0178302.

Zhou T, Chou J, Watkins PB, Kaufmann WK. 2009. Toxicogenomics: transcription profiling for toxicology assessment. Exs 99:325-366.

Zhou YH, Cichocki JA, Soldatow VY, Scholl EH, Gallins PJ, Jima D, et al. 2017. Editor's Highlight: Comparative Dose-Response Analysis of Liver and Kidney Transcriptomic Effects of Trichloroethylene and Tetrachloroethylene in B6C3F1 Mouse. Toxicol Sci 160:95-110.