NTP
National Toxicology Program

Implementation of Systematic Review by the National Toxicology Program Office of Health Assessment and Translation (OHAT)

Kristina Thayer, Ph.D. Director, OHAT

Division of the NTP
National Institute of Environmental Health Sciences

OpenTox USA
29 - 30 October 2013
Hamner Conference Center, North Carolina Biotechnology Center,
Raleigh-Durham, USA

Outline

- Overview of systematic review and OHAT framework for integrating evidence
- Assessing study quality and utility
- Data management and display software tools

Systematic Review

- A scientific investigation that focuses on a specific question, and uses explicit, pre-specified methods to identify, select, summarize, and assess the findings of similar studies
- Provides greater transparency
- Used to:
  - Reach evidence-based conclusions
  - Clarify need for additional research
  - May or may not result in quantitative meta-analysis
- Existing methodologies are primarily used for assessment of healthcare interventions
  - e.g., Cochrane, AHRQ, GRADE

What Does A Systematic Review Not Do?

- Does not eliminate the need for expert judgment
- Improves but does not guarantee reproducibility of conclusions across separate evaluations
  - Increased transparency does not necessarily eliminate differences in scientific judgment
- Existing methods do not provide guidance on how to
  - Reach hazard identification conclusions
  - Integrate evidence across human, animal, and mechanistic studies

OHAT Systematic Review and Evidence Integration Framework

Step 1: Prepare topic

Step 2: Search for and select studies

Step 3: Extract data from studies

Step 4: Assess individual study quality

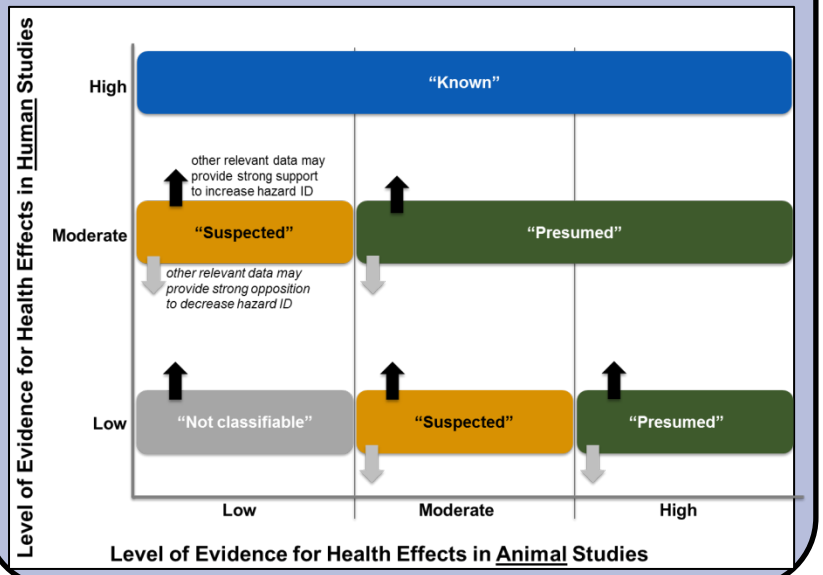
Step 5: Rate confidence in body of evidence (animal and human studies)

Initial Confidence by Key Features of Study Design	Factors Decreasing Confidence	Factors Increasing Confidence	Confidence in the Body of Evidence
High (++++) 4 Features	❖ Risk of Bias ❖ Unexplained Inconsistency ❖ Indirectness ❖ Imprecision ❖ Publication Bias	❖ Large Magnitude of Effect ❖ Dose Response ❖ All Plausible Confounding • Studies report an effect and residual confounding is toward null • Studies report no effect and residual confounding is away from null ❖ Consistency • Across animal models or species • Across dissimilar populations • Across study design types ❖ Other e.g., particularly rare outcomes	High (++++)
Moderate (+++) 3 Features			Moderate (+++)
Low (++) 2 Features			Low (++)
Very Low (+) ≤1 Features			Very Low (+)

Features
- Controlled exposure
- Exposure prior to outcome
- Individual outcome data
- Comparison group used

Step 6: Translate confidence ratings into level of evidence for health effect

Step 7: Integrate evidence to develop hazard identification conclusions



Evaluating evidence for biological plausibility provided by *in vitro*, cellular, genomic, or mode of action data

Strong Support **Weak Support**

- ***Relevance of biological process or pathway to human health***
- ***Consistency***
- ***Relevance of concentration***
- ***Potency***
- ***Dose response***
- ***Study quality***
- ***Publication bias***

Factors considered parallel elements used to evaluate confidence in the other data streams

A conclusion of "strong" support may lead to a label of "suspected" in the absence of human or animal data

Evaluating Study Quality and Utility

Definitions: Study Quality and Utility

- **Reporting quality:** How well was the study reported?
- **Internal validity or risk of bias:** How credible are the findings based on design and conduct of the study?
  - Reporting quality checklist ≠ risk of bias tool
    - Non-reporting has negative impact on risk of bias (attempts will be made to follow up with study authors)
  - Single summary scores of studies strongly discouraged, domain-based approach recommended
- **Directness and applicability:** How well does the study address the topic under review?
  - Route, timing/duration of exposure and health outcome assessment
  - Upstream indicator of primary health outcome?
  - Relevance of animal model for human health

Same set of questions applied to different study designs

Risk of Bias Domain	Criterion	Animal	Controlled Exposure	Cohort	Case-Control	Cross-sectional	Case Series
Selection	Was administered dose or exposure level adequately randomized?	X	X				
	Was allocation to study groups adequately concealed?	X	X				
	Were the comparison groups appropriate?			X	X	X	
Confounding	Did the study design or analysis account for important confounding and modifying variables?	X	X	X	X	X	X
	Did researchers adjust or control for other exposures that are anticipated to bias results?	X	X	X	X	X	X
Performance	Were experimental conditions identical across study groups?	X	X				
	Did deviations from the study protocol impact the results?	X	X	X	X	X	X
	Were the research personnel and human subjects blinded to the study group during the study?	X	X				
Attrition	Were outcome data incomplete due to attrition or exclusion from analysis?	X	X	X	X	X	
Detection	Were the outcome assessors blinded to study group or exposure level?	X	X	X	X	X	X
	Were confounding variables assessed consist measures						
	Can we be confident in the exposure charact						
	Can we be confident in the outcome assessment?	X	X	X	X	X	X
Reporting	Were all measured outcomes reported?	X	X	X	X	X	X
Other	Were there any other potential threats to internal validity (e.g., inappropriate statistical methods)?	X	X	X	X	X	X

Some items seem unlikely to be useful in short-term but may be useful in long-term, i.e., changes in reporting quality, develop empirical data to assess potential risk of bias of item

Current Tool: Response Format & Review Process

- Uses responses recommended by the Clarity Group
  - “definitely no” (●) risk of bias
  - “probably no” (●) risk of bias
  - “probably yes” (●) risk of bias (rule for non-reported elements)
  - “definitely yes” (●) risk of bias
- Rationale for selecting a response is noted
  - Based on instructions and expert judgment (e.g., members of review team, technical advisors)
- Risk of bias is independently assessed by 2 members of review team
  - Independent reviews discussed to develop draft response for report
- Risk of bias conclusions assessed by review team, technical advisors, and undergo external public peer-review

Risk of Bias Ratings Across Individual Studies

++ Definitely Low risk of bias
+ Probably Low risk of bias
- Probably High risk of bias
-- Definitely High risk of bias
○ Not applicable due to study design

Draft OHAT Risk of Bias Questions	NotGood ., 2010	Bucher et al., 1999	Wolfe et al., 2000	Boyles et al., 2011	Thayer et al., 2008
Selection Bias					
Was administered dose or exposure level adequately randomized?	○	○	○	○	○
Was allocation to study groups adequately concealed?	○	○	○	○	○
Were the comparison groups appropriate?	+	++	+	--	+
Confounding Bias					
Did the study design or analysis account for important confounding and modifying variables?	-	++	-	--	+
Did researchers adjust or control for other exposures that are anticipated to bias results?	-	++	-	-	--
Performance Bias					
Were experimental conditions identical across study groups?	○	○	○	○	○
Did deviations from the study protocol impact the results?	-	++	+	-	-
Were the research personnel and human subjects blinded to the study group during the study?	○	○	○	○	○
Attrition / Exclusion Bias					
Were outcome data incomplete due to attrition or exclusion from analysis?	--	++	-	+	+
Information / Detection Bias					
Were outcome assessors blinded to study group or exposure group?	+	++	+	+	+
Were confounding variables assessed consistently across groups using valid and reliable measures?	--	++	+	++	++
Can we be confident in the exposure characterization?	-	++	--	-	+
Can we be confident in the outcome assessment?	-	++	-	+	-
Selective Reporting Bias					
Were all measured outcomes reported?	+	++	+	-	+

Study Quality and Utility are Assessed in Several Different Steps

Step 1: Prepare topic

Step 2: Search for and select studies

Step 3: Extract data from studies

Step 4: Assess individual study quality (risk of bias)

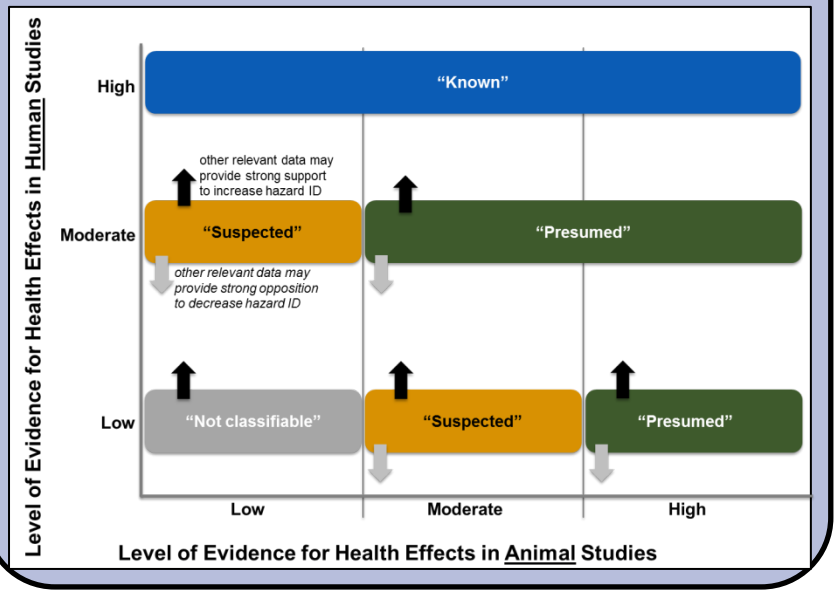
Step 5: Rate confidence in body of evidence (animal and human studies)

Initial Confidence by Key Features of Study Design	Factors Decreasing Confidence	Factors Increasing Confidence	Confidence in the Body of Evidence
High (++++) 4 Features	❖ Risk of Bias ❖ Unexplained Inconsistency ❖ Indirectness ❖ Imprecision ❖ Publication Bias	❖ Large Magnitude of Effect ❖ Dose Response ❖ All Plausible Confounding • Studies report an effect and residual confounding is toward null • Studies report no effect and residual confounding is away from null ❖ Consistency • Across animal models or species • Across dissimilar populations • Across study design types ❖ Other e.g., particularly rare outcomes	High (++++)
Moderate (+++) 3 Features			Moderate (+++)
Low (++) 2 Features			Low (++)
Very Low (+) ≤1 Features			Very Low (+)

Features
- Controlled exposure
- Exposure prior to outcome
- Individual outcome data
- Comparison group used

Step 6: Translate confidence ratings into level of evidence for health effect

Step 7: Integrate evidence to develop hazard identification conclusions



Software Tools to Manage Systematic Review, Data Collection, & Visual Display

Data Management Goals

- Create data repositories of extracted data
  - Reduce duplication of effort – data extraction currently done by people = $$$$$
- Disseminate data files of extracted data to public
  - Increase transparency
  - Facilitate independent and supplemental analysis of data
  - Data would reside in CEBS (and potentially other places)
- Create more concise reports
  - Use figures and tables to summarize studies
  - Web-based reports
  - Interactive modules (e.g. HAWC)

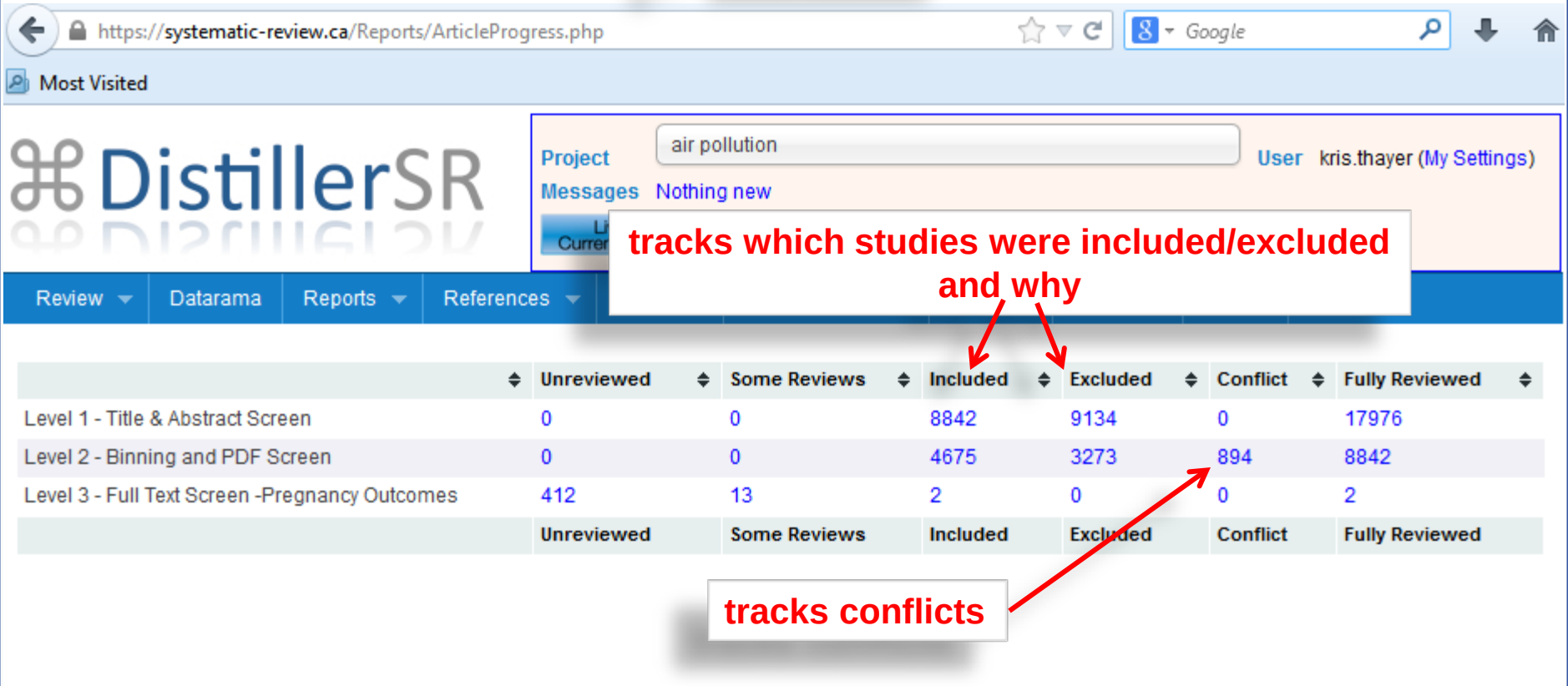
Data Management & Display Tools

- Distiller Systematic Review® Software = systematic review software to help manage screening process
- ICF International's DRAGON for data extraction and warehousing
- HAWC (Health Assessment Workspace Collaborative) for visual display and other analysis functions
  - Ivan Rusyn's talk in session C on Thursday

Distiller Systematic Review® Software

Distiller SR®

web-based



upload references from Endnote or other reference managers
– PDFs of articles can also be uploaded

Distiller SR® - Customizable Screening Level Forms

Y. Kim and B. K. Lee. 2011. Association between urinary arsenic and diabetes mellitus in the Korean general population according to KNHANES 2008. Sci Total Environ 409(19): 4054-62.

Submit Form and go to This Form - Next Reference or Skip to Next

INTRODUCTION: We present data from the Korean National Health and Nutrition Examination Survey (KNHANES) 2008 on the associations between urinary arsenic and diabetes mellitus in a representative sample of the adult Korean population.

METHODS: This study was based on data obtained in KNHANES 2008, which was conducted for three years (2007-2009) using a rolling sampling design involving a complex, stratified, multistage, probability-cluster survey of a representative sample of the noninstitutionalized civilian population of South Korea.

RESULTS: Geometric means of total urinary arsenic concentration in females and total participants with diabetes mellitus were significantly higher than in participants without diabetes mellitus after adjustment for covariates, including age, seafood consumption, body mass index (BMI), hypertension, area of residence, regional area, education level, and smoking and drinking status. Multiple regression analysis after similar adjustment showed that total urinary arsenic concentration was associated with diabetes status in the females and total participants. In addition, after similar adjustment, the odds ratios (ORs) for diabetes mellitus in female participants and all participants were 1.502 (95% CI, 1.038-2.171) and 1.312 (95% CI, 1.040-1.655), respectively, for doubling of the level of urinary total arsenic concentration.

CONCLUSION: This study showed an association between total urinary arsenic concentration and the prevalence of diabetes mellitus in a representative sample of the adult population, especially women, with environmental arsenic exposure after adjustment for seafood intake and relevant diabetes risk factors.

Do the title or abstract suggest the article contains original data related to a topic of interest?

- ○ yes
- ○ yes (non-English)
- ○ yes, but is a review, commentary, or letter with no original data
- ○ no, not relevant
- ○ not directly relevant, but could be supportive material
- ○ unsure

Add text here to described screening level criteria for relevance:

Inclusion criteria

- bullet format
- bullet format

Exclusion criteria

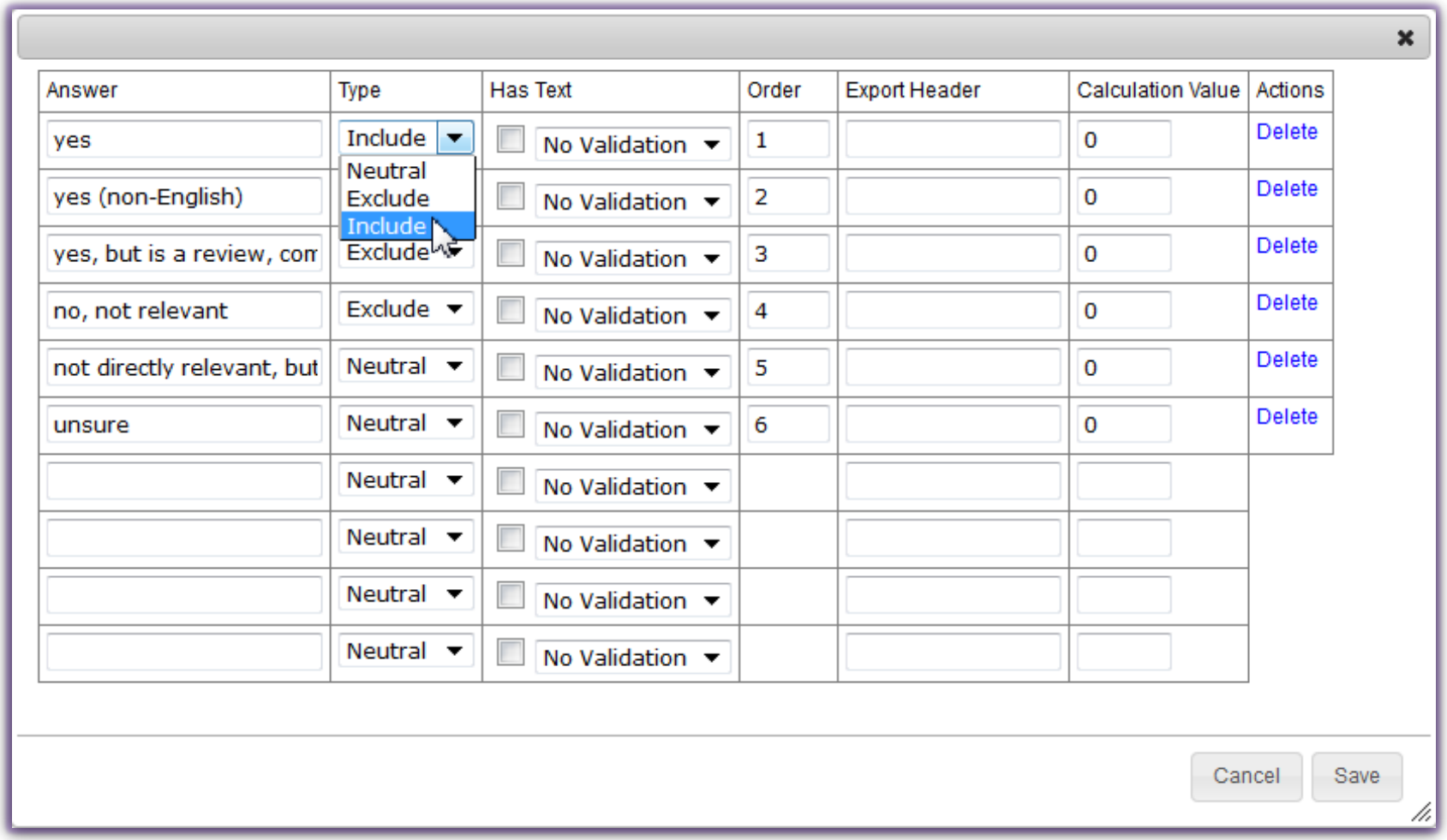
- bullet format
- bullet format

Comments

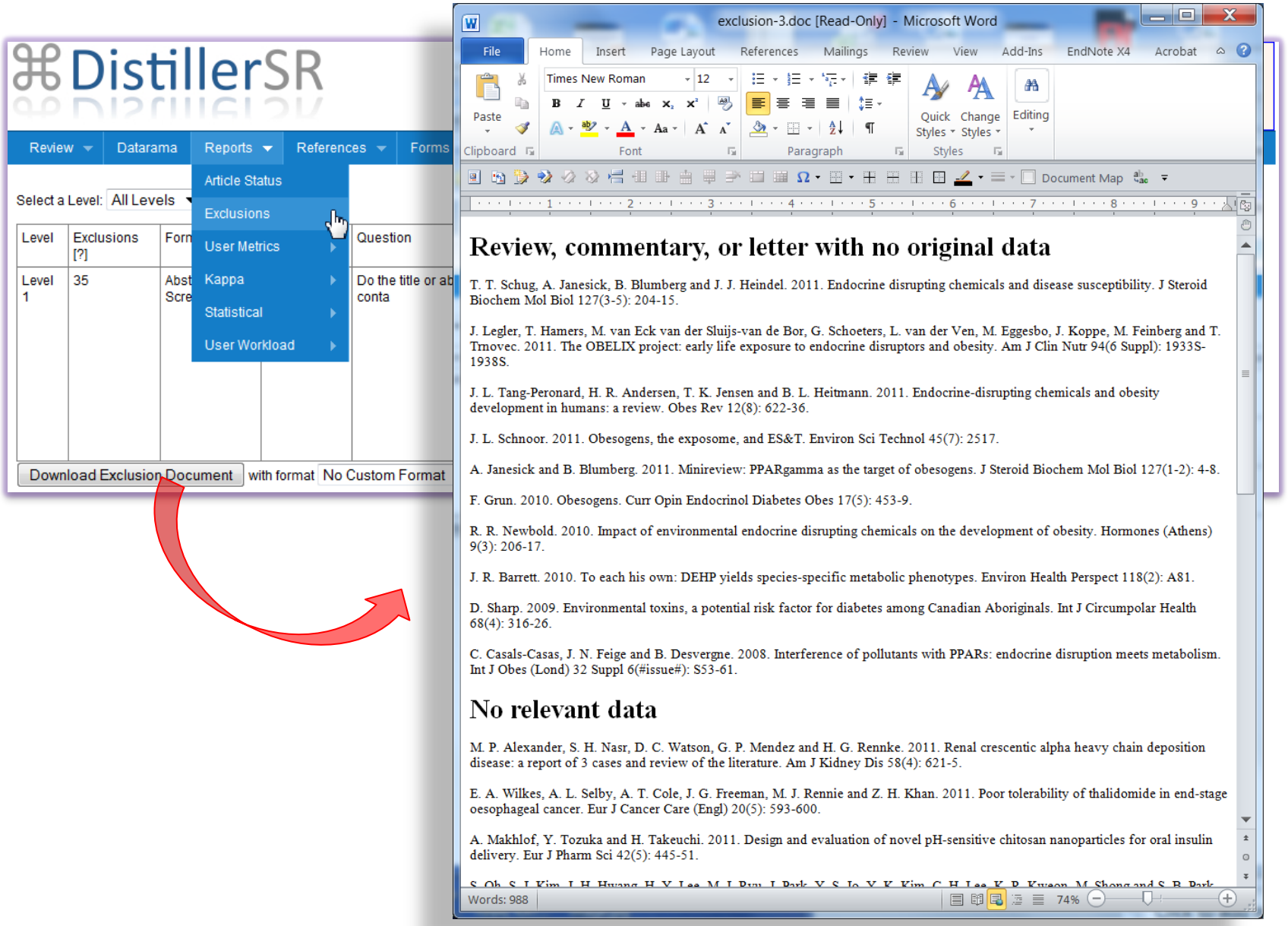
Submit Form and go to This Form - Next Reference or Skip to Next

View Audit Log

Distiller SR® – Inclusion & Exclusion Settings



Distiller SR® - Exclusion Reports



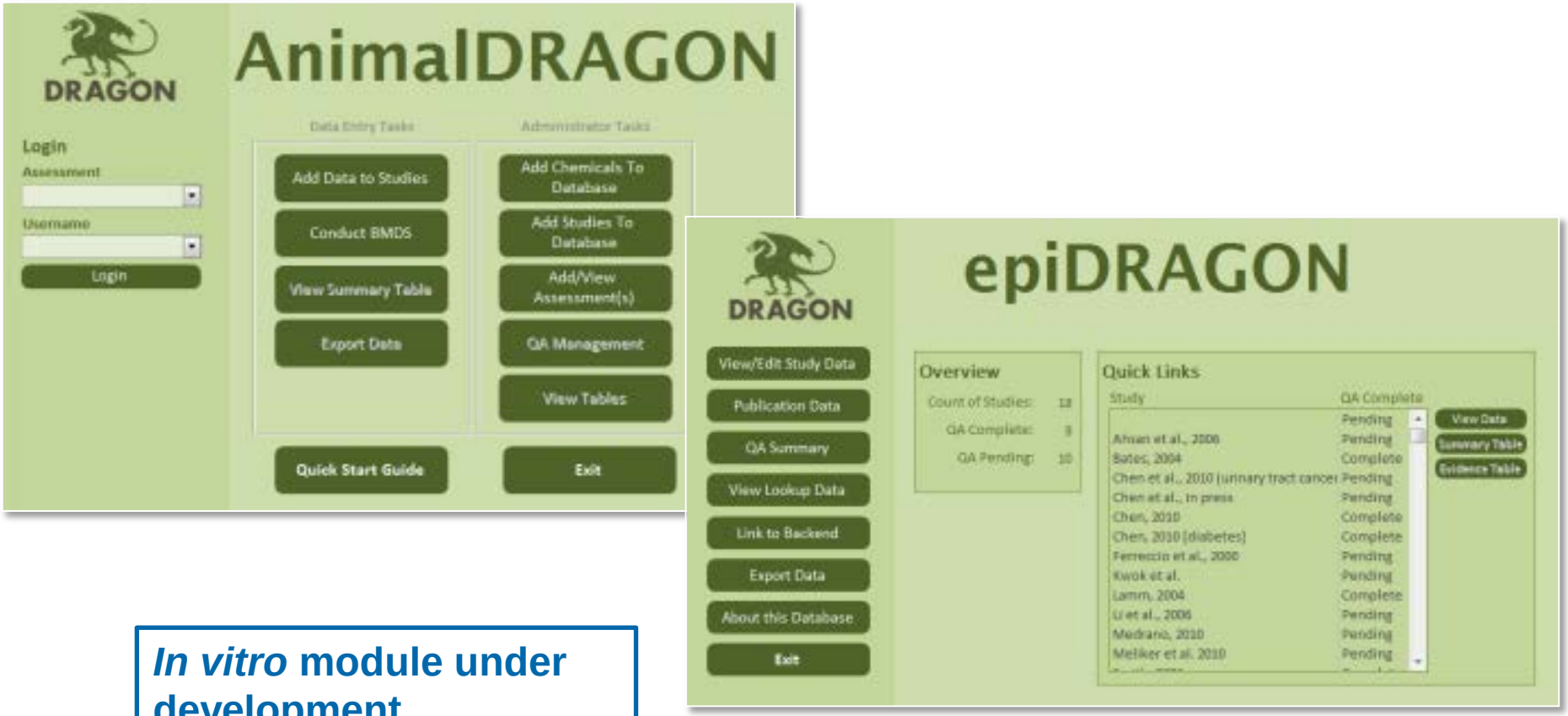
Storing Data and Decisions during Systematic Review

The ICF DRAGON Development Team
DRAGON@icfi.com
Audrey Turley (919-293-1621)

Team members: Cara Henning, AJ Overton, Kristen Martin, Pam Ross, Joshua Cleland, Audrey Turley

DRAGON

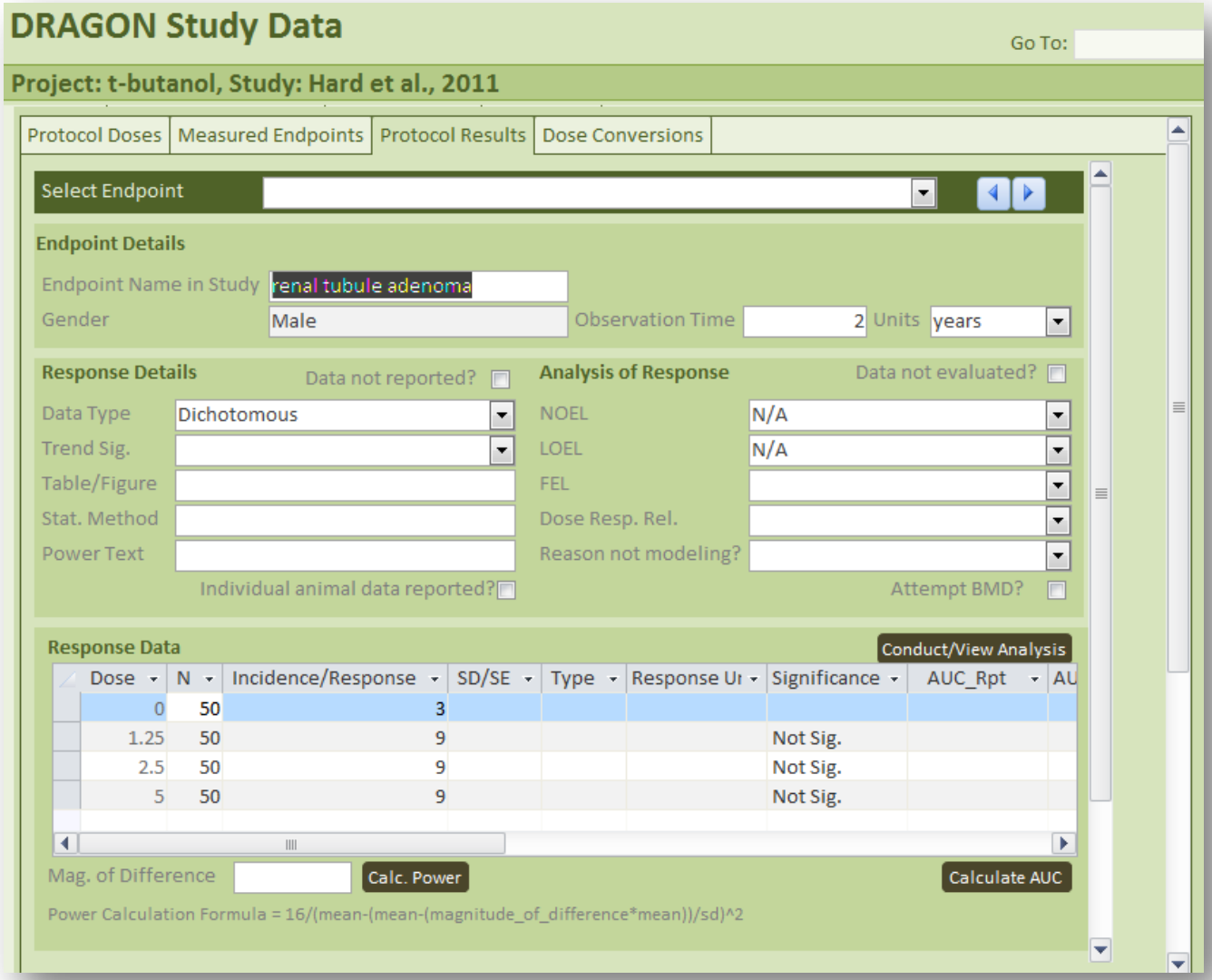
Data Entry to Facilitate Creation of Evidence Tables



In vitro **module under development**

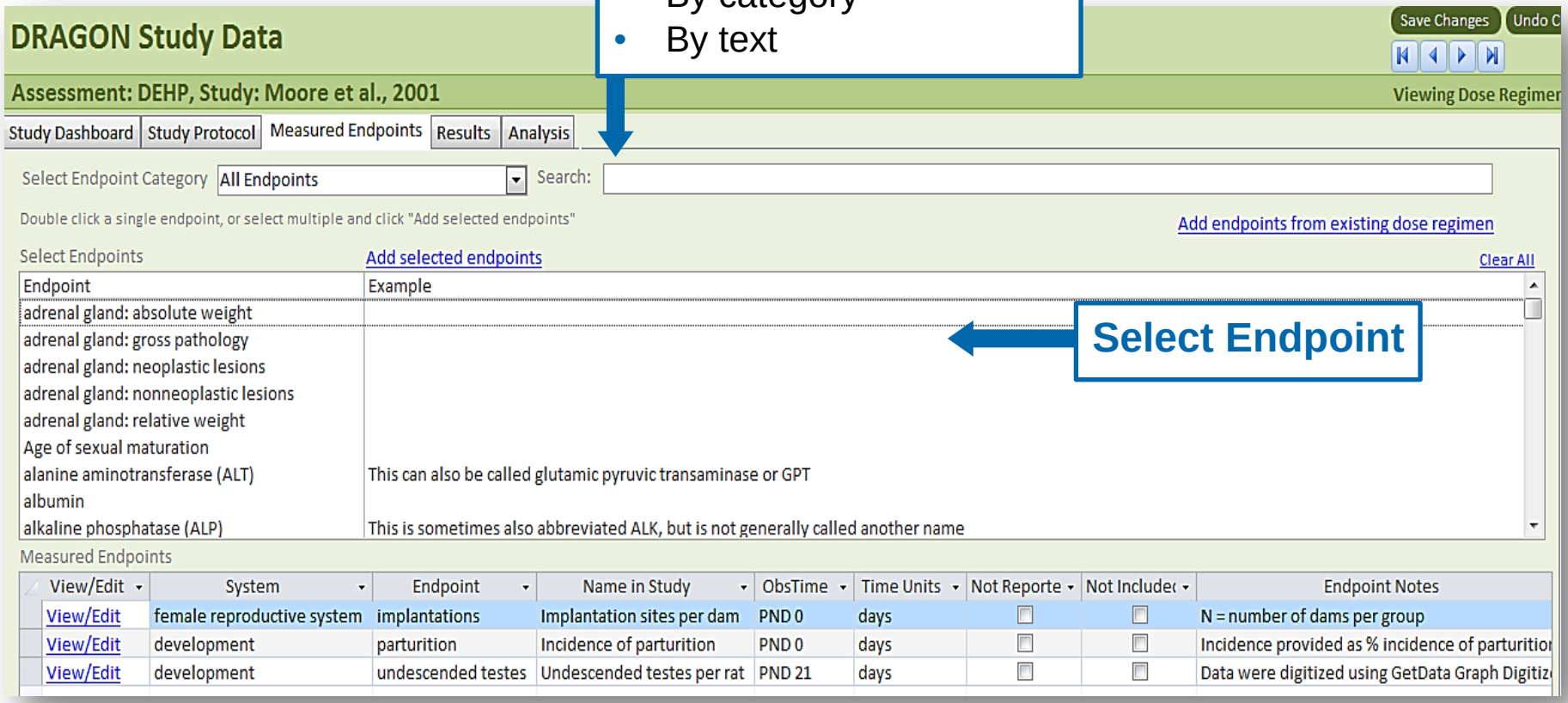
animalDRAGON

animalDRAGON
Animal Data Extraction



Manage Endpoints

Search for Endpoints
- By category
- By text

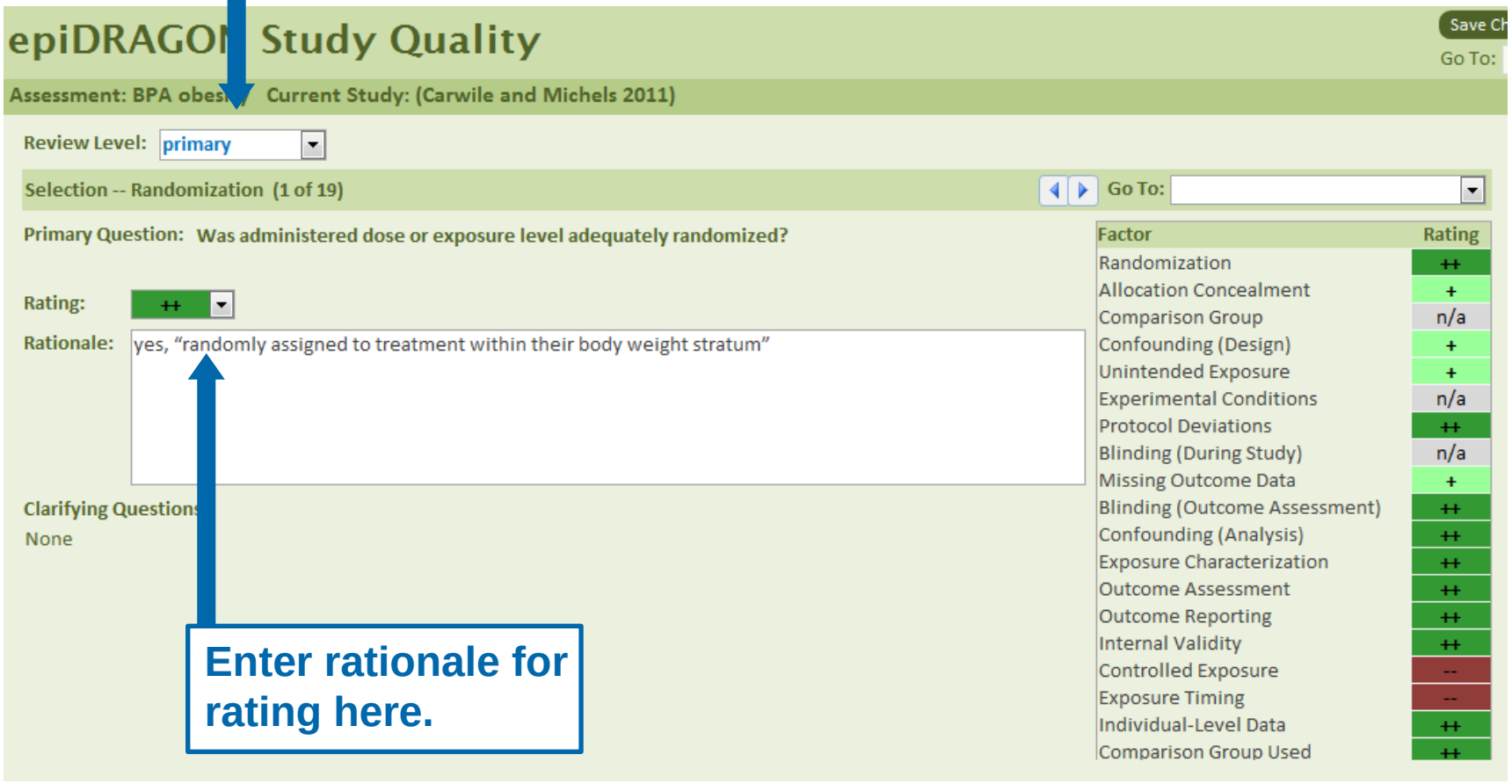


Select Endpoint

Add Details
- Name in study
- Notes
- View/Edit more

Study Quality / Risk of Bias

Quality is assessed by two separate reviewers.



Enter rationale for rating here.

Study Quality / Risk of Bias: Management Console

Quality Control

epiDRAGON QA Management

- QA Status: Pending

Short Citation	Data Group for QA				
	Study and Popluation	Study Effects	Exposure Information	Statistical Findings	Senior Review
	☐	☐	☐	☐	☐
Ahsan et al., 2006	☐	☐	☐	☐	☐
Bates, 2004	☑	☑	☑	☑	☑
Chen et al., 2010 (urinary tract cancer)	☐	☐	☐	☐	☐
Chen et al., in press	☐	☐	☐	☐	☐
Chen, 2010	☑	☑	☑	☑	☑
Chen, 2010 [diabetes]	☑	☑	☑	☑	☑
Ferreccio et al., 2000	☐	☐	☐	☐	☐
Kwok et al.	☐	☐	☐	☐	☐
Lamm, 2004	☑	☑	☑	☑	☑
Li et al., 2006	☐	☐	☐	☐	☐
Medrano, 2010	☐	☐	☐	☐	☐
Meliker et al. 2010	☐	☐	☐	☐	☐
Smith, 2006	☑	☑	☑	☑	☑
Tseng, 1996	☑	☑	☑	☑	☑
von Ehrenstein, 2006	☑	☑	☑	☑	☑
Wu et. al	☐	☐	☐	☐	☐
Xia, 2009	☑	☑	☑	☑	☑

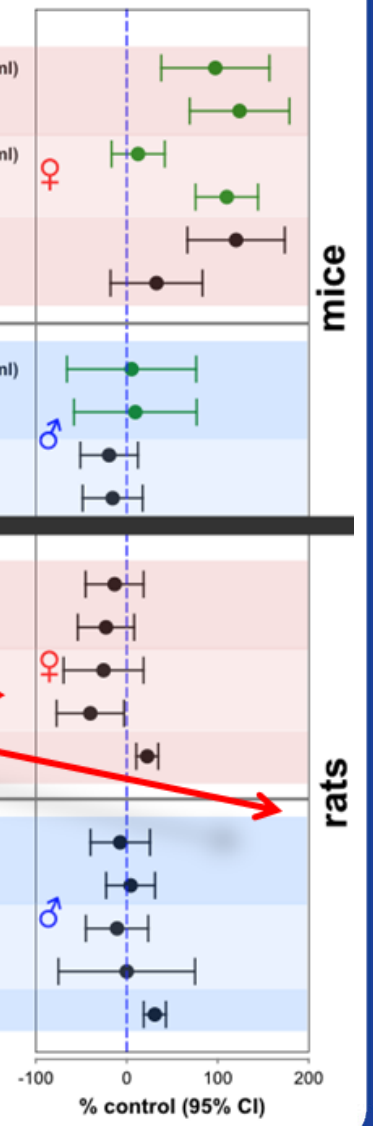
Interactive Visual Display Tools to Help Evaluate Complex Data Sets (MetaDataViewer/HAWC)

Animal Data

Reference	Species, Strain (sex, generation, n)	Treatment Period (age at assessment)	Route	Dose (mg/kg/d)	Endpoint
Alonso-Magdalena et al. 2010	mouse, OF-1 (♀ P 6)	GD9-GD16 (4m post-partum)	sc injection, dam	0.01	leptin, plasma (fasted, ng/ml)
				0.1	
	mouse, OF			0.01	leptin, plasma (fasted, ng/ml)
				0.1	
Miyawaki et al. 2007	mouse, ICR (♀ F1 16-19)	GD10-PND30 (PND30)	drinking water, dam	0.26	leptin, serum [ng/ml]
				2.72	
Alonso-Magdalena et al. 2010	mouse, OF-1 (♂ F1 6-13)	GD9-GD16 (6m)	sc injection, dam	0.01	leptin, plasma (fasted, ng/ml)
				0.1	
Miyawaki et al. 2007	mouse, ICR (♂ F1 19-25)	GD10-PND30 (PND30)	drinking water, dam	0.26	leptin, serum [ng/ml]
				2.72	
Ferguson et al. 2011	rat, Sprage Dawley (♀ F1 13-15)	GD6-21, PND1-21 to pup (PND21)	drinking water	0.0025	ghrelin, serum (ng/ml)
				0.025	
Ferguson et al. 2011	rat, Sprage Dawley (♀ F1 13-15)	GD6-21, PND1-21 to pup (PND21)	drinking water	0.0025	leptin, serum (ng/ml)
Wei et al. 2011	rat, Wistar (♀ F1 6)	GD0-			(ng/ml)
Ferguson et al. 2011	rat, Sprage Dawley (♂ F1 10-17)	GD6-21, PND1-21 to pup (PND21)	drinking water	0.0025	ghrelin, serum (ng/ml)
				0.025	
Ferguson et al. 2011	rat, Sprage Dawley (♂ F1 10-17)	GD6-21, PND1-21 to pup (PND21)	drinking water	0.0025	leptin, serum (ng/ml)
				0.025	
Wei et al. 2011	rat, Wistar (♂ F1 6)	GD0-PND21 (27w)	gavage, dam	0.05	leptin, serum (ng/ml)

sort by up to 2 excel columns
(e.g., reference & endpoint)

visually stratify by up to 2 variables
(e.g., species & sex)



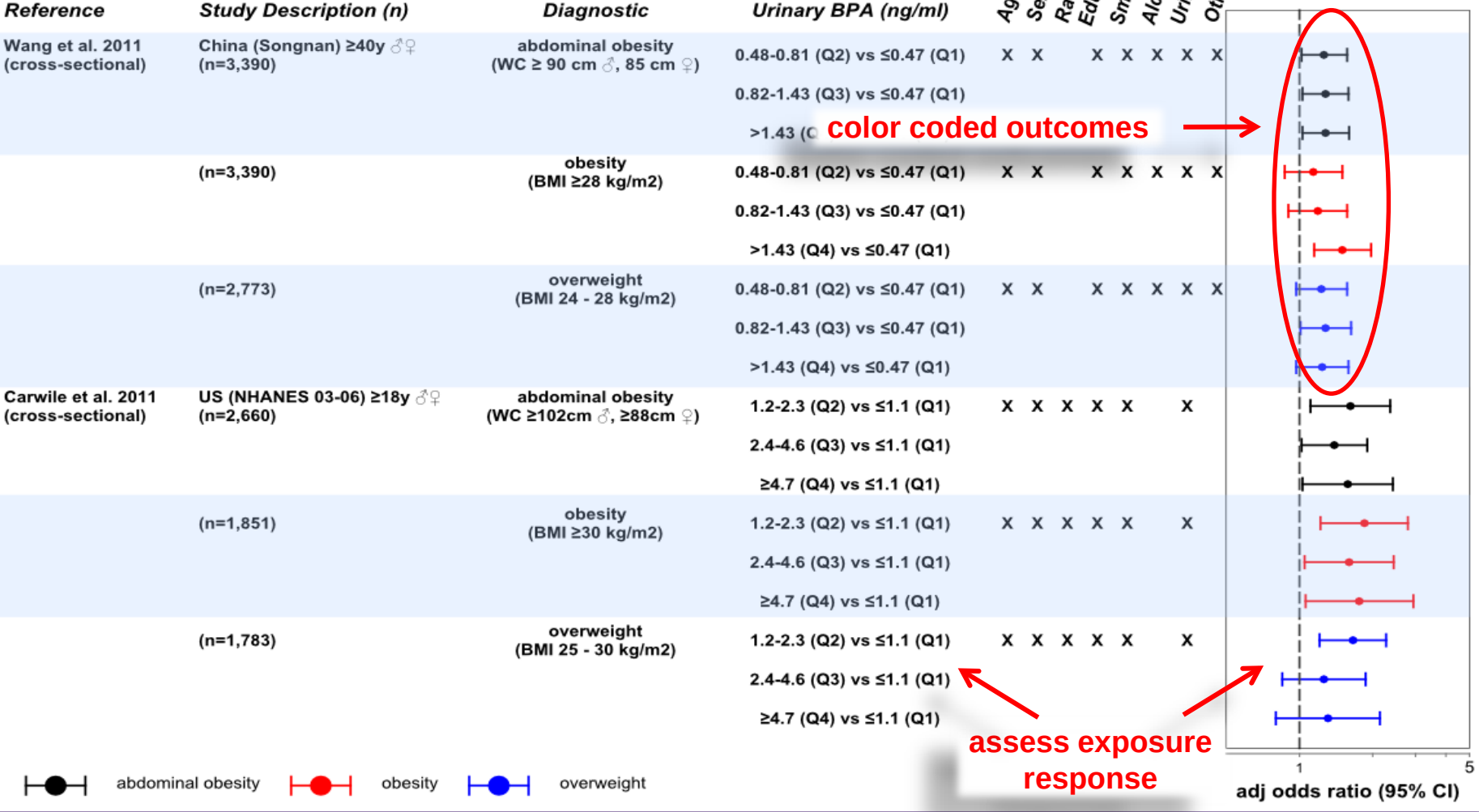
oral
sc injection
♀ female
♂ male

Human Data

Reference	Study Description (n)	Diagnostic	Urinary BPA (ng/ml)	Age	Sex	Race	Education	Smoking	Alcohol	Urine Cr	Other
Wang et al. 2011 (cross-sectional)	China (Songnan) ≥40y ♂♀ (n=3,390)	abdominal obesity (WC ≥ 90 cm ♂, 85 cm ♀)	0.48-0.81 (Q2) vs ≤0.47 (Q1)	X	X		X	X	X	X	X
			0.82-1.43 (Q3) vs ≤0.47 (Q1)								
			>1.43 (Q								
	(n=3,390)	obesity (BMI ≥28 kg/m2)	0.48-0.81 (Q2) vs ≤0.47 (Q1)	X	X		X	X	X	X	X
			0.82-1.43 (Q3) vs ≤0.47 (Q1)								
			>1.43 (Q4) vs ≤0.47 (Q1)								
	(n=2,773)	overweight (BMI 24 - 28 kg/m2)	0.48-0.81 (Q2) vs ≤0.47 (Q1)	X	X		X	X	X	X	X
			0.82-1.43 (Q3) vs ≤0.47 (Q1)								
			>1.43 (Q4) vs ≤0.47 (Q1)								
Carwile et al. 2011 (cross-sectional)	US (NHANES 03-06) ≥18y ♂♀ (n=2,660)	abdominal obesity (WC ≥102cm ♂, ≥88cm ♀)	1.2-2.3 (Q2) vs ≤1.1 (Q1)	X	X	X	X	X		X	
			2.4-4.6 (Q3) vs ≤1.1 (Q1)								
			≥4.7 (Q4) vs ≤1.1 (Q1)								
	(n=1,851)	obesity (BMI ≥30 kg/m2)	1.2-2.3 (Q2) vs ≤1.1 (Q1)	X	X	X	X	X		X	
			2.4-4.6 (Q3) vs ≤1.1 (Q1)								
			≥4.7 (Q4) vs ≤1.1 (Q1)								
	(n=1,783)	overweight (BMI 25 - 30 kg/m2)	1.2-2.3 (Q2) vs ≤1.1 (Q1)	X	X	X	X	X		X	
			2.4-4.6 (Q3) vs ≤1.1 (Q1)								
			≥4.7 (Q4) vs ≤1.1 (Q1)								

color coded outcomes

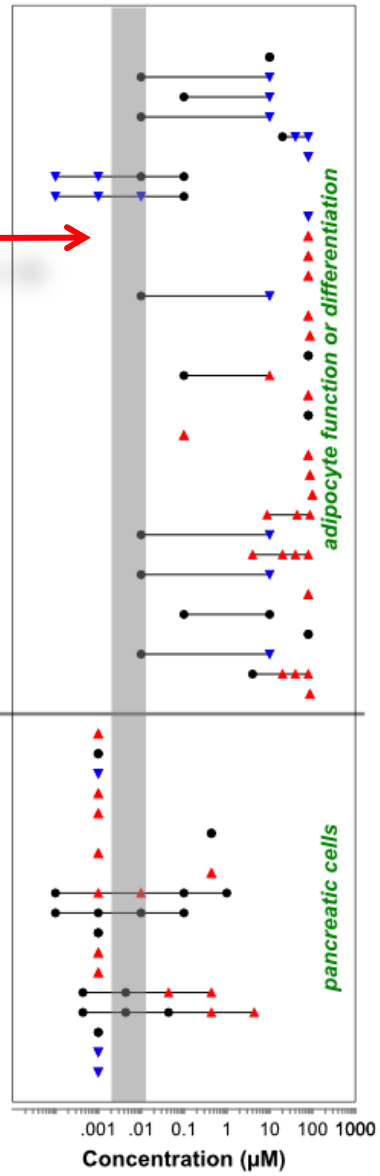
assess exposure response



In Vitro Data

Reference	*Cell Type*	*Concentrations, μM**	*Endpoint*
Biemann 2011	C3H/10T1/2 MSC	10	adipocyte number
Biemann 2011	C3H/10T1/2 MSC	0.010, 10*	adipocyte number
Lee 2008	3T3-L1 fibroblast	0.1, 10*	adiponectin mRNA
Biemann 2011	C3H/10T1/2 MSC	0.010, 10*	adiponectin mRNA
Kidani 2010	3T3-L1 fibroblast	20, 40*, 80*	adiponectin production
Kidani 2010	3T3-L1 fibroblast	80*	adiponectin protein
Hugo 2008	adipose (human breast ♀)	0.0001*, 0.001*, 0.01, 0.1	adiponectin release
Hugo 2008	adipose (human abdominal ♀)	0.0001*, 0.001*, 0.01*, 0.1	adiponectin release
Kidani 2010	3T3-L1 fibroblast	80*	
Masuno 2005	3T3-L1 fibroblast	80*	
Masuno 2005	3T3-L1 fibroblast	80*	
Phrakonkham 2008	3T3-L1 fibroblast	80*	
Biemann 2011	C3H/10T1/2 MSC	0.010, 10*	FABP4 mRNA
Phrakonkham 2008	3T3-L1 fibroblast	80*	FAS mRNA
Masuno 2002	3T3-L1 fibroblast	[87.6*]	GPDH activity
Phrakonkham 2008	3T3-L1 fibroblast	80	leptin concentration
Lee 2008	3T3-L1 fibroblast	0.1, 10*	leptin mRNA
Phrakonkham 2008	3T3-L1 fibroblast	80*	leptin mRNA
Phrakonkham 2008	3T3-L1 fibroblast	80	lipid accumulation
Sargis 2010	3T3-L1 fibroblast	0.1*	lipid accumulation
Masuno 2005	3T3-L1 fibroblast	80*	lipid accumulation
Masuno 2002	3T3-L1 fibroblast	[87.6*]	lipid accumulation
Wang 2010	3T3-L1 fibroblast	100*	lipid accumulation
Masuno 2002	3T3-L1 fibroblast	[8.76*, 43.8*, 87.6*]	LPL activity
Biemann 2011	C3H/10T1/2 MSC	0.010, 10*	LPL mRNA
Masuno 2005	3T3-L1 fibroblast	80*	LPL mRNA
Biemann 2011	C3H/10T1/2 MSC	0.010, 10*	PPARγ2 mRNA
Phrakonkham 2008	3T3-L1 fibroblast	80*	Pref-1 mRNA
Lee 2008	3T3-L1 fibroblast	0.1, 10	resistin mRNA
Phrakonkham 2008	3T3-L1 fibroblast	80	triglyceride content
Biemann 2011	C3H/10T1/2 MSC	0.010, 10*	triglyceride content
Masuno 2005	3T3-L1 fibroblast	4, 20*, 40*, 80*	triglyceride content
Masuno 2002	3T3-L1 fibroblast	[87.6*]	triglyceride content
Quesada 2002	islets, β-cells (Swiss OF1 mouse)	0.001*	[Ca2+]i concentration
Soriano 2012	islets, β-cells (ERβ -/-, C57 mouse)	0.001	[Ca2+]i oscillation frequency
Alonso-Magdalena 2005	islets, α-cells (Swiss OF1 mouse)	0.001*	[Ca2+]i oscillation frequency
Nadal 2000	islets, β-cells (Swiss OF1 mouse)	0.001*	[Ca2+]i oscillation frequency
Soriano 2012	islets, β-cells (ERβ WT, C57 mouse)	0.001*	[Ca2+]i oscillation frequency
Makaji 2011	TC-6 β cells	[0.438]	CHOP/GADD153 protein
Quesada 2002	islets, β-cells (Swiss OF1 mouse)	0.001*	CREB phosphorylation
Makaji 2011	TC-6 β cells	[0.438*]	HSP70, GRP78, GRP94 proteins
Alonso-Magdalena 2008	islets (Swiss OF1 mouse)	0.0001, 0.001*, 0.01*, 0.1, 1	insulin content
Soriano 2012	islets (C57 mouse)	0.0001, 0.001, 0.01, 0.1	insulin release
Soriano 2012	islets (ERβ -/-, C57 mouse)	0.001	insulin secretion
Soriano 2012	islets (ERβ WT, C57 mouse)	0.001*	insulin secretion
Soriano 2012	islets (human)	0.001*	insulin secretion
Adachi 2005	islets (Wistar rat)	[0.000438, 0.00438, 0.0438*, 0.438*]	insulin secretion
Makaji 2011	TC-6 β cells	[0.000438, 0.00438, 0.0438, 0.438*, 4.38*]	insulin secretion
Soriano 2012	islets (ERβ -/-, C57 mouse)	0.001	KATP channel activity
Soriano 2012	islets (ERβ WT, C57 mouse)	0.001*	KATP channel activity
Soriano 2012	islets, β-cells (human)	0.001*	KATP channel activity

indicate physiologically relevant concentration range



● no statistical significant effects observed

▲ statistically significant increase

▼ statistically significant decrease

estimated range of unconjugated BPA in human blood samples (Vandenberg et al. 2010)

Current Status & Next Steps

- See http://ntp.niehs.nih.gov/go/38673 for documents and updates
- Draft OHAT Approach (posted February 2013)
  - Currently applying approach to 2 case-study protocols
    - BPA exposure and obesity
    - PFOA/PFOS exposure and immunotoxicity
  - Host “lessons learned” webinar (Spring/Summer 2014)
- “Beta testing” DRAGON and HAWC
- Working to improve transparency on how we consider other types of data, e.g., *in vitro*, HTS

Acknowledgements

- **Office of Health Assessment and Translation**
  - Abee Boyles
  - Kembra Howdeshell
  - Andrew Rooney, Deputy Director
  - Michael Shelby
  - Kyla Taylor
  - Kristina Thayer, Director
  - Vickie Walker
- **Office of Liaison, Policy and Review**
  - Mary Wolfe, Director
  - Lori White
- **Technical Advisors and Experts**
  - **Lisa Bero,** Director, San Francisco Branch, United States Cochrane Center at UC San Francisco
  - **Gordon Guyatt,** Co-chair, GRADE Working Group, McMaster University
  - **Malcolm Macleod,** CAMARADES Centre, University of Edinburgh
  - **Karen Robinson,** Co-Director, Evidence-Based Practice Center, The Johns Hopkins Bloomberg School of Public Health
  - **Holger Schünemann**, Co-chair, GRADE Working Group, McMaster University
  - **Tracey Woodruff,** Director, Program on Reproductive Health and the Environment, UCSF
- **NTP BSC Working Group**
  - **Lynn Goldman, Chair,** Dean, School of Public Health and Health Services, George Washington University, Washington, DC
  - **Reeder Sams, Vice-chair,** Acting Deputy Director, National Center for Environmental Assessment/RTP Division, USEPA
  - **Lisa Bero,** Director, San Francisco Branch, United States Cochrane Center at UC San Francisco
  - **Edward Carney,** Senior Science Leader, Mammalian Toxicology, Dow Chemical Company
  - **David Dorman,** Professor, North Carolina State University
  - **Elaine Faustman**, Director, Institute for Risk Analysis and Risk Communication, University of Washington
  - **Dale Hattis**, Research Professor, George Perkins Marsh Institute, Clark University
  - **Malcolm Macleod,** CAMARADES Centre, University of Edinburgh
  - **Tracey Woodruff,** Director, Program on Reproductive Health and the Environment, UCSF
  - **Lauren Zeise,** Chief, Reproductive and Cancer Hazard Assessment Branch, OEHHA, California EPA