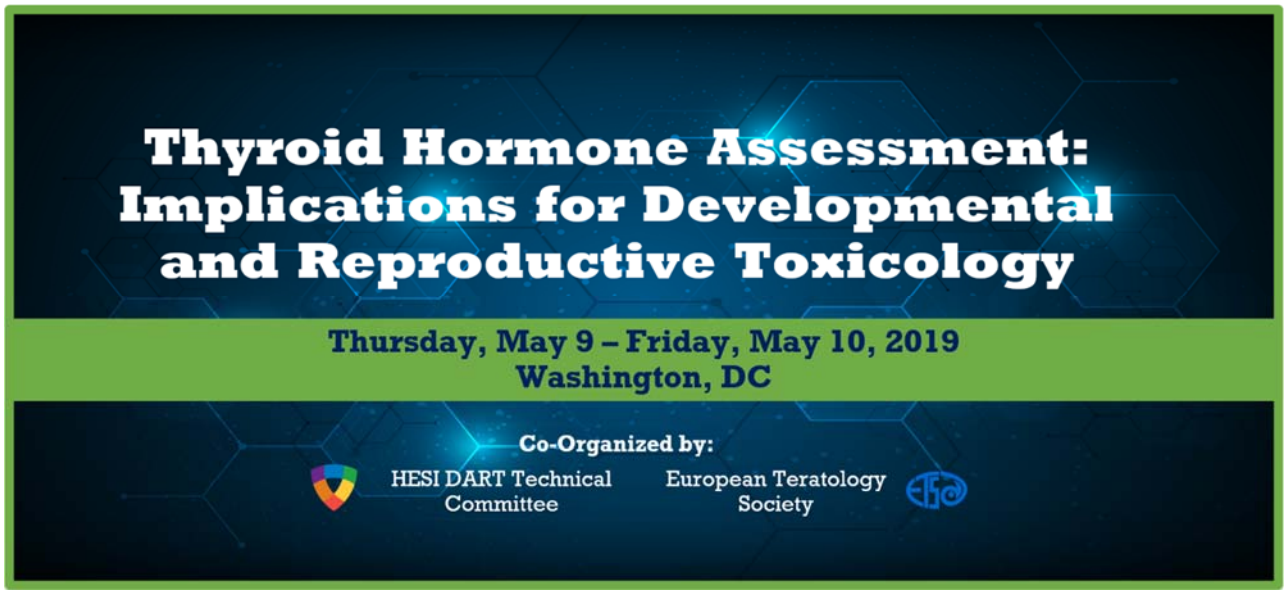

Recent updates to OECD developmental/reproductive toxicology guidelines and other regulatory guidelines and guidance require the measurement of thyroid hormone levels in the blood of mammalian laboratory species during development. Preliminary analyses indicate that there is a wide variability across laboratories in the methods being used to measure thyroid hormones in young rodents, as well as in the success of obtaining reliable data. Even though publicly available regulatory guidelines and guidance address study design, they allow varied approaches to thyroid hormone measurement in rodents, and an optimal study design or logical approach to thyroid hormone testing in young rodents has not yet been established in a regulatory testing context. Validity, accuracy, sensitivity and reproducibility of the assays are issues of concern. It is not clear to what extent variability in the data can be attributed to methodological issues or to innate biological variability.

This workshop aims to:

(1) Present the state-of-the-science on thyroid hormone assessments, specifically as it relates to preclinical methods and data collection, and identify gaps and knowledge as it relates to regulatory DART testing,
(2) Provide clarification and guidance regarding the collection (timing and methods), assessment (standardization and validation), and interpretation of thyroid hormone data (as it relates to adversity) for regulatory toxicology and risk assessment,
(3) Discuss and come to consensus on recommendations on how to improve data interpretation/understanding of thyroid changes and their relationship to adverse outcomes

AGENDA

DAY 1, THURSDAY MAY 9, 2019

Session 1: Setting the stage

8:30am – 8:40am	**Welcome & workshop charge** *Susan Makris, US EPA*
8:40am – 9:10am	**The Thyroid Axis – Overview of Anatomy, Physiology, Regulation in Mammalian Systems** *Mary Gilbert, US EPA*
9:10am – 9:40am	**An AOP Network for Thyroid Hormone Disruption and Adverse Outcomes** *Kevin Crofton, R3Fellows, LLC*
9:40am – 10:10am	**Mild thyroid dysfunction during pregnancy; consequences for pregnancy outcome and fetal development** *Robin Peeters, Erasmus University*
10:10am – 10:30am	**Break**

Session 2: Methodologies & Interpretation

10:30am – 11:00am	**Regulatory Requirements for Evaluation of Thyroid Status** *Sue Marty, Dow*
11:00am – 11:30am	**Pathology Endpoints: Evaluation of Thyrotoxicants in Animal Studies** *Brent Walling, Charles River*
11:30am – 12:00pm	**Thyroid hormone perturbation and neurodevelopmental toxicity in animal studies** *Ellen Hessel, RIVM*
12:00pm – 12:30pm	**Overview of in vitro and non-mammalian assays to investigate chemicals for thyroid-axis disrupting potential** *Michael Hornung, US EPA*
12:30pm – 1:15pm	**Lunch**

Session 3: Global thyroid activities

1:15pm – 1:30pm	**Overview of ongoing thyroid activities in the European Union and United States** *Manon Beekhuijzen, Charles River*
1:30pm – 1:45pm	**Overview of BfR laboratory survey** *Olga Kucheryavenko, German Federal Institute for Risk Assessment (BfR)*
1:45pm – 2:15pm	**HESI DART- ETS Thyroid Hormone Survey – Results & the Path Forward** *Pragati Coder, Charles River*

Session 4: Breakout Groups

2:15pm – 2:30pm	Breakout group instructions/**break**
2:30pm – 4:30pm	Groups are charged with identifying key data gaps and research needs, recommendations and future directions as it relates to: 1) Specimen collection, assay methodologies, sample analysis and reporting 2) Data interpretation, regulatory aspects and risk assessment
4:30pm – 4:45pm	Wrap-up
5:00pm – 6:30pm	**Light Reception**

DAY 2, FRIDAY MAY 10, 2019

8:30am – 8:35am **Day 1 Recap and Charge for Day 2**

Session 5: Break-out group report-back

8:35am – 10:30am **Breaking groups will report back on their discussions.**

10:30am – 10:50am **Break**

Session 6: Perspectives on thyroid testing to improving interpretation

This session will present regional perspectives from both regulators and the regulated. Regulators will provide an overview of their agency's expectations for thyroid testing and outline critical issues regarding interpretation. Regulated industry panelists will address the critical issues encountered during study testing and when results are brought to the regulators and discuss their strategies for responding to regulators.

10:50am – 11:15am European Perspectives

- *Niklas Andersson, ECHA*
- *Nina Hallmark, Bayer*

Asian Perspectives

- Hiroaki Aoyama, Institute for Environmental Toxicology (on behalf of Japanese Food Safety Commission)
- Tomoya Yamada, Sumitomo Chemical

North American Perspectives

- Elizabeth Mendez, US EPA
- Bethany Hannas, Corteva
- Jennifer Foreman, ExxonMobil
- Miyun Tsai-Turton, US FDA
- LaRonda Morford, Lilly

11:15 – 12:15pm **Panel discussion**

Facilitators: Susan Makris, US EPA and Pragati Coder, Charles River

Session 7: Looking ahead

12:15pm – 12:45pm **Identification of research needs and recommendations: A facilitated discussion**

Facilitators: Aldert Piersma (RIVM), Alan Hoberman (Charles River)

12:45pm – 1:00pm **Closing remarks & next steps**

Susan Makris, US EPA and Pragati Coder, Charles River

BREAKOUT GROUP – LIST OF QUESTIONS & PROPOSED GROUPING

PRIMARY THEME – THYROID ENDPOINTS ON REPRO. & DEV. TOX. STUDIES.

Shared Goal: key recommendations, data gaps, future directions

- *3 groups per group topic; each group will have different question prioritized so all questions will be covered*
- *Prepare summary sheet of baseline information/assumptions as a starting point.*
- *Breakout group will receive pared down questions; faciliators will have more detailed*

	GROUP 1 **SPECIMEN COLLECTION, ASSAY METHODOLOGIES, SAMPLE ANALYSIS AND REPORTING**	**GROUP 2** **DATA INTERPRETATION, REGULATORY ASPECTS AND RISK ASSESSMENT**
The Charge:	*Identify and prioritize critical aspects of specimen collection, sample analysis and reporting.*	*Identify key scientific criteria influencing data interpretation, regulatory decisions and risk assessment.*
Data Gaps	Do you agree with the identified data gaps from speakers? Priorities? Omissions from the data gaps? What over and above the guidelines should be required (thyroid and non-thyroid endpoints)? How do you assess critical windows of susceptibility? Besides collection of thyroid data (serum hormone concentrations, organ weight and histopathology), what other endpoints would/could provide context or a basis for interpretation? (e.g., brain weights, liver weights and/or histopathology, pituitary?). Would you recommend that these endpoints be included in study guidelines put forth by health and regulatory authorities? (Question: Would you include similar endpoints also for all EATS endpoints? Or just for the Thyroid?) In your experience, what other information has been useful in interpreting thyroid/endocrine data?	Do you agree with the identified data gaps from speakers? Priorities? Omissions from the data gaps? What are the critical endpoints (thyroid and non-thyroid endpoints) that should be collected and aren't being collected? What types of research do we need to understand critical windows of susceptibility? - Reverse T3 is universally accepted to be a hormone with no known biological effect? Do you think there is value in this endpoint with relevance to regulatory decisions? - What additional data is informative when making decisions regarding true biological effects? Do endpoints such as brain, pituitary or liver weights and/or histopathology of these organs provide context or relevant evidence to aid in interpretation? In your experience, at what point does a change in thyroid endpoints – hormone, organ weight, histopathology, become adverse? How does one assess thyroid homeostasis? Furthermore, is there a

	Neurobehavioral endpoints?	threshold/CV/reference range beyond which one would consider the results biologically significant? OECD 407 and OECD 408 Guidelines provide minimum acceptability criteria for variability (%CV). Should these criteria be included in other guidelines put forth by health and regulatory authorities?
Variability	What are the key factors influencing variability? Animal-to-Animal? Collection Methodologies? Husbandry (e.g., how do you minimize stress)? Assay Methodologies and Criteria? How can variability be minimized?	What are the key factors influencing data interpretation? How does animal-to-animal variability affect data interpretation? How can variability be minimized (i.e., husbandry)?
Basis for Interpretation, acceptance criteria	Do you agree with the minimum acceptability/validation criteria set forth by the speakers and based on the survey data? Are there additional criteria that should be included in a guidance document? If so, what should they be? (*Complicated feedback loop mechanism and whether or not it's a permanent change/reversibility*) OECD 407 and OECD 408 Guidelines provide minimum acceptability criteria for variability (%CV). Should these criteria be included in other guidelines put forth by health and regulatory authorities – why or why not?	What kind of statistical analyses should be recommended for thyroid (hormone) data? And for other endpoints? How does one differentiate true changes from statistical anomalies? OECD 421 and 422 Guidelines require only the assessment of T4 (and other hormones, if warranted). How do significant changes in one hormone influence data interpretation decisions and should additional hormones be included in these TGs to allow for a more thorough assessment of endocrine effects? (Note: OECD TG 443 requires only T4 and TSH and OECD TG 414 requires T3, T4, and TSH). Would you recommend that the full suite of thyroid hormones be included in study guidelines put forth by health and regulatory authorities? Would you recommend that free T3 and reverse T3 be included in the same battery of tests – why or why not?
Reference Compounds?	What reference compounds are used routinely in your laboratory for positive control (e.g., 6-PTU, methimazole etc.) and for negative control (besides the vehicle control)? Would you recommend that a list of reference compounds be included in study guidelines put forth by health and	Does reference compound data provide useful information to assist in overall data interpretation? If so, how? And if so, would you recommend that a list of reference compounds be included in study guidelines put forth by health and regulatory authorities? (Note: moderator to generate list of

	regulatory authorities? (Note: moderator to generate list of reference compounds that have been demonstrated as positive or negative modulators of the HPT axis).	reference compounds that have been demonstrated as positive or negative modulators of the HPT axis).
Human Relevance	What are reasonable recommendations regarding preclinical testing in the context of human relevance? What is the right animal model (knock-in, knock-out)? Are there informative alternative assays?	Placing the animal data in context of the clinic - how do you apply the lessons learned from reference ranges set in humans? Is there standardized methodology that can be used to compare animal data with human data, or extrapolation?
Public Database?	Do you see benefit of a publicly-available historical control (reference) database? Would you, or your company, be willing to routinely contribute data to said database on an annual or biannual basis to help improve its value over time? Should such a database be publicly-available, would you recommend additional endpoints/variables that were not included in the HESI Survey data?	Do you see benefit of a publicly-available historical control (reference) database? Would such a database assist in interpretation of your data, and if so, how? How can you optimally use the data, considering strength and limitations of the data? Is there additional biological context that can be used to supplement the raw data?
		How do you (or do you) interpret the correlation between liver and thyroid hormone? How do you interpret offspring effects? What has been most informative to companies running these assays? Has there been something informative to companies?