OECD Extended Advisory Group for Molecular Screening and Toxicogenomics (EAGMST)

Guidance Document for Consistent Reporting of 'Omics Data From Various Sources

Transcriptomics Reporting Framework (TRF)

EAGMST Meeting, Paris FR
June 29th, 2018

Disclaimer

- *The views expressed in this presentation are those of the author(s) and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.*

Agenda

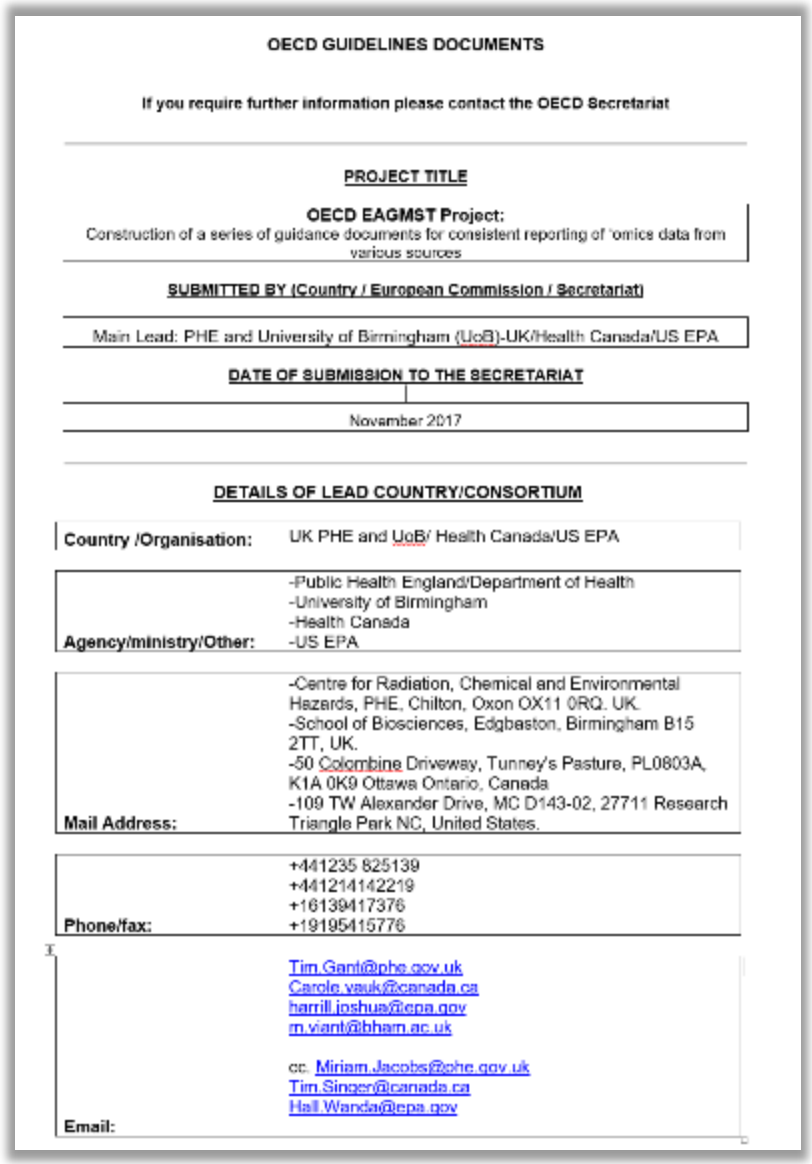
- **Background**
  - Regulatory Acceptance of transcriptomic data
  - ‘Omics Reporting Frameworks
- **Transcriptomics Reporting Framework (TRF)**
  - Objective & Scope
  - Topic Areas
  - Document Structure
  - Case Studies

Workgroups

  - Membership & Leadership
  - Charge
- **Timelines**
  - Overall Project Timeline
  - Progress-to-Date
- **Q & A**

Reasons for lack of regulatory acceptance of transcriptomic data

1) Poor experimental design and data quality (early days) = bad reputation for omics technologies
2) Lack of accepted quality control standards and data quality assessment tools
3) Lack of availability of metadata necessary for interpretation and regulatory application
4) Lack of transparency, public availability and best practices/standards for data processing methods
5) Variances in methods and prior knowledge used to analyse and interpret genomics data
6) Lack of standardized reporting frameworks to ensure that all required and appropriate data, metadata, and analytical processes are available.



OECD GUIDELINES DOCUMENTS

If you require further information please contact the OECD Secretariat

PROJECT TITLE

OECD EAGMST Project:
Construction of a series of guidance documents for consistent reporting of 'omics data from various sources

SUBMITTED BY (Country / European Commission / Secretariat)

Main Lead: PHE and University of Birmingham (UoB)-UK/Health Canada/US EPA

DATE OF SUBMISSION TO THE SECRETARIAT

November 2017

DETAILS OF LEAD COUNTRY/CONSORTIUM

Country /Organisation:	UK PHE and UoB/ Health Canada/US EPA
Agency/ministry/Other:	-Public Health England/Department of Health -University of Birmingham -Health Canada -US EPA
Mail Address:	-Centre for Radiation, Chemical and Environmental Hazards, PHE, Chilton, Oxon OX11 0RQ. UK. -School of Biosciences, Edgbaston, Birmingham B15 2TT, UK. -50 Colombine Driveway, Tunney's Pasture, PL0803A, K1A 0K9 Ottawa Ontario, Canada -109 TW Alexander Drive, MC D143-02, 27711 Research Triangle Park NC, United States.
Phone/fax:	+441235 825139 +441214142219 +16139417376 +19195415776
Email:	Tim.Gant@phe.gov.uk Carole.vauk@canada.ca harrill.joshua@epa.gov m.viant@bham.ac.uk cc. Miriam.Jacobs@phe.gov.uk Tim.Singer@canada.ca Hall.Wanda@epa.gov

ecetoc
EUROPEAN CENTRE FOR ECOTOXICOLOGY AND TOXICOLOGY OF CHEMICALS

Buesen et al. Reg Tox Pharm. 2017

Project Description

To develop frameworks for the standardisation of reporting of 'omics data generation and analysis, to ensure that all of the information required to understand, interpret and reproduce an 'omics experiment and its results are available.

Purpose: to ensure that sufficient information is available to enable an evaluation of the quality of the experimental data and interpretation, and support reproducibility.

NOT to stipulate the methods of data analysis or interpretation….**Rather**, provide guidance on reporting of information that fosters transparency and reproducibility.

Project Name	Project Lead
Metabolomics Reporting Framework (MRF)	Mark Viant (U. Birmingham, UK)
Transcriptomics Reporting Framework (TRF)	Joshua Harrill (USEPA) Carole Yauk (Health Canada)
Reference Baseline Analysis (RBA)	Tim Gant (PHE, UK)

TRF Objective, Working Group Charge & Scope

OBJECTIVE: Development of a Transcriptomics Reporting Framework (TRF) for processing of 'omics data that will facilitate acceptance of transcriptomics studies in a regulatory setting.

WORKING GROUP CHARGE: The TRF working group is tasked with determining what information should be captured by the TRF to support interpretation and computational reproducibility of 'omics experiments by members of the regulatory community. Such information will also be of value to researchers in academia and industry.

SCOPE: The transcriptomics reporting framework (TRF) is a tool for documenting the details of laboratory-based toxicology studies that utilize a transcriptomics technology: i.e. an assay that measures the abundance of many transcripts simultaneously and that provides highly multiplexed outputs. The TRF is appropriate for use in documenting experiments involving the use of either *in vivo* or *in vitro* laboratory models. The information captured by the TRF should be of sufficient detail for other researchers to replicate all aspects of the transcriptomics experiment including administration of chemicals, sample processing, raw data collection and computational methods used to generate processed data. The TRF is designed to be coupled with downstream analysis reporting modules (DARMs) that detail the steps and resources necessary to reproduce a computational analysis of the processed data. Specific DARMs are coupled to the TRF based on the researcher's specific use case.

TRF Document, Major Topic Areas

EXPERIMENT:

- The experiment should be described in sufficient detail that would allow another researcher to replicate the experiment.
- Adapted from existing sources
- Information in this section is independent of ‘omics platform

PROCESSING AND ANALYSIS OF ‘OMICS DATA:

- The transcriptomics technology, sample processing procedures, methods used to collect raw data and methods used to generate processed data.
- Described in Gant et al. (2017).
- Information in this section is dependent on ‘omics platform

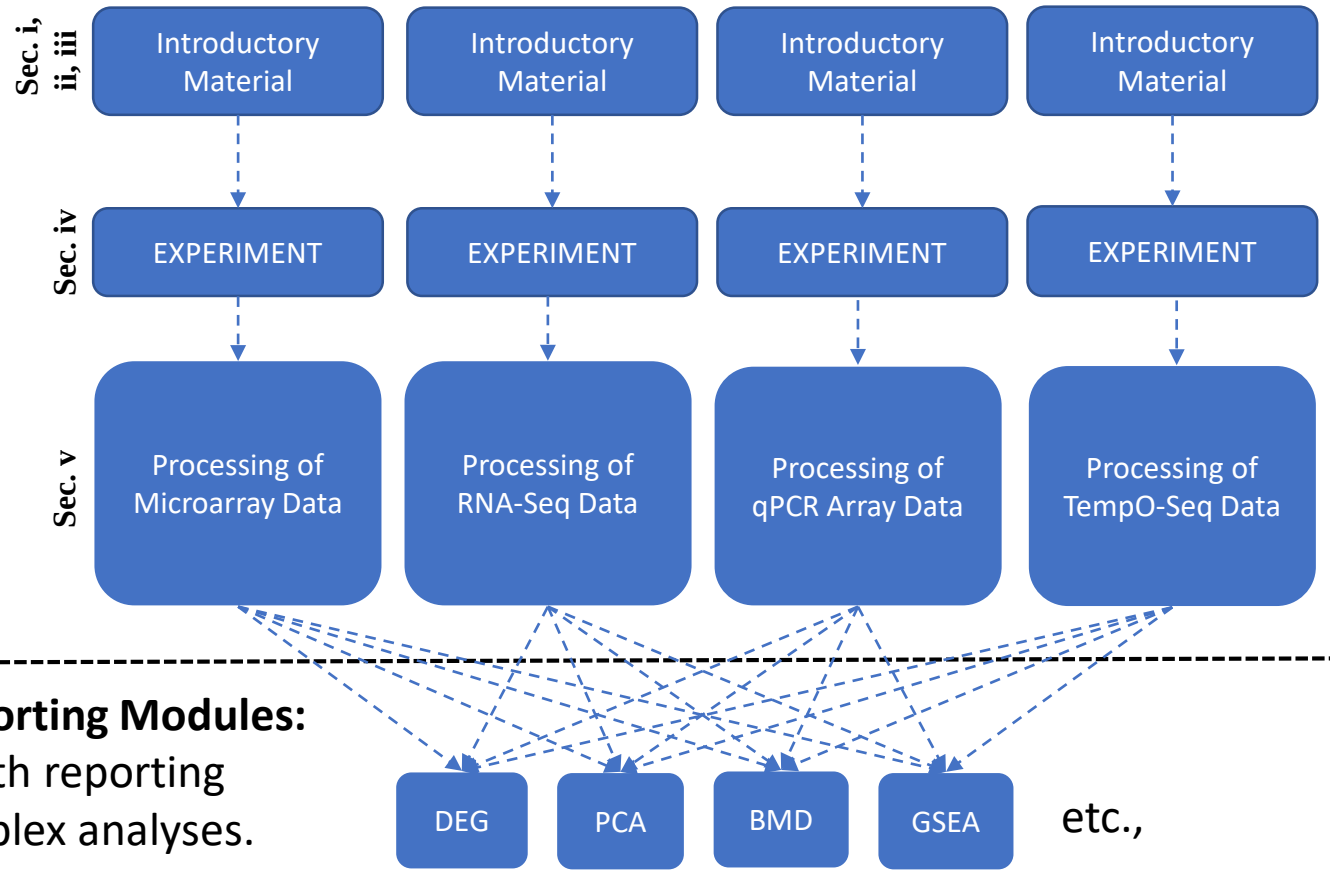
DOWNSTREAM ANALYSIS REPORTING MODULES [DARMs]

- Detail the steps and resources necessary to reproduce a computational analysis of the processed data.

TRF Document Structure

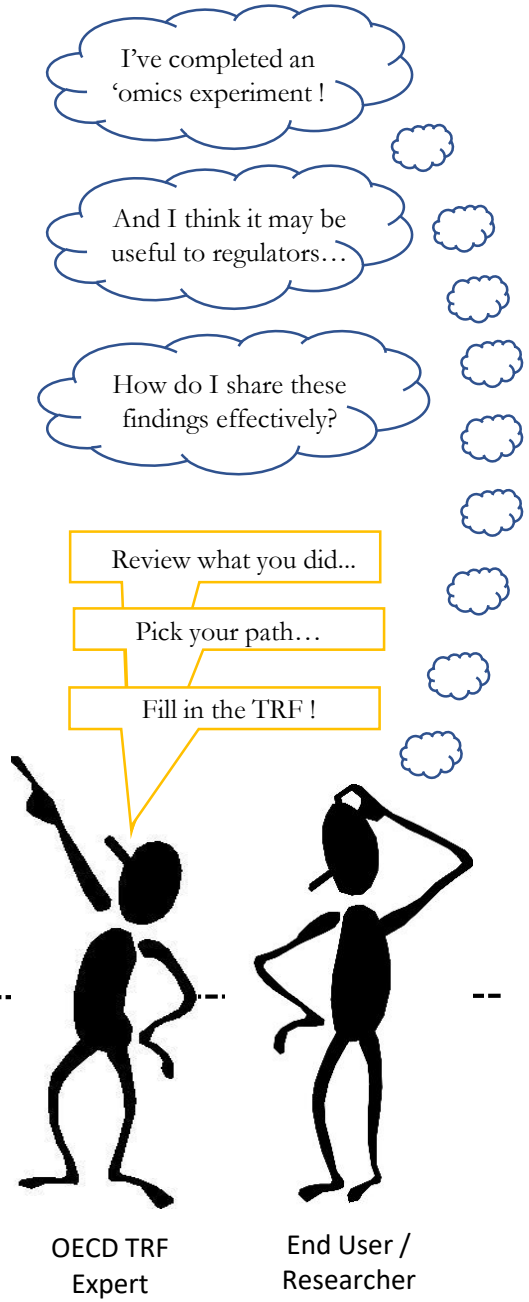
Transcriptomics Reporting Framework (TRF)

- Several technology-specific documents
- Redundancy in sections i-iv across documents



Downstream Analysis Reporting Modules:

- Separate documents with reporting structure for more complex analyses.
- To be mixed and matched with TRF based on use case.



Draft TRF Outline

i. ABSTRACT

ii. INTRODUCTION

I. Purpose / Aims
II. Background
III. Scope
IV. Related 'Omics Standard Projects

iii. DEFINITIONS / ABBREVIATIONS

iv. EXPERIMENT

I. Study Rationale
II. Study Design
III. Subject / Test System Characteristics
IV. Test Article
V. Treatment Conditions
VI. Study Exit
VII. Sample Collection & Pre-processing
VIII. Sample Identification Codes
IX. Supporting Data Streams

iv. PROCESSING OF 'OMICS DATA

I. Technology
II. Sample Processing
III. Transcriptomics Study Design
IV. Specification of Raw Data
V. Data Normalization
VI. Data filtering
VII. Identification and Removal of Low Quality or Outlying Datasets

- **Stylistic alignment:**
  - Previous OECD guidance in the biological sciences (where applicable)
  - Metabolomics Reporting Framework (MRF) – *In Progress*
- **Reporting Format**
  - Narrative text followed by Reporting Fields
- **Consistent vocabulary across modules**
- **Database compatibility (?)**

TRF Document Outline, Introductory Material

i. **ABSTRACT**

ii. **INTRODUCTION**
   I. Purpose / Aims
   II. Background
   III. Scope
   IV. Related ‘Omics Standard Projects

iii. **(TABLE OF) DEFINITIONS / ABBREVIATIONS**

These sections will be drafted by the leadership team and sent to the entire project group for comment.

TRF Document Outline, Experiment

iv. EXPERIMENT

I. Study Rationale
II. Study Design
III. Subject / Test System Characteristics
IV. Test Article
V. Treatment Conditions
VI. Study Exit
VII. Sample Collection & Pre-processing
VIII. Sample Identification Codes
IX. Supporting Data Streams

Content is technology <u>independent</u>

Section to be drafted by a **section workgroup** under guidance of a **section leader**

Content leverages previously existing works

Experiment Module, Existing Resources

CEBS

- Waters, M., et al., (2003). Systems toxicology and the Chemical Effects in Biological Systems (CEBS) knowledge base. *EHP Toxicogenomics, 111*(1T), 15-28. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12735106
- Fostel, J. M., et al., (2007). Toward a checklist for exchange and interpretation of data from a toxicology study. *Toxicol Sci, 99*(1), 26-34. doi:10.1093/toxsci/kfm090

MIAME

- Brazma, A., et al., (2001). Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet, 29*(4), 365-371. doi:10.1038/ng1201-365

ToxRTool

- Schneider, K., et al., (2009). "ToxRTool", a new tool to assess the reliability of toxicological data. *Toxicol Lett, 189*(2), 138-144. doi:10.1016/j.toxlet.2009.05.013
- Segal, D., et al., (2015). Evaluation of the ToxRTool's ability to rate the reliability of toxicological data for human health hazard assessments. *Regul Toxicol Pharmacol, 72*(1), 94-101. doi:10.1016/j.yrtph.2015.03.005

SOAR

- McConnell, E. R., et al., (2014). Systematic Omics Analysis Review (SOAR) tool to support risk assessment. *PLoS One, 9*(12), e110379. doi:10.1371/journal.pone.0110379

TRF Document Outline, ‘OMICS DATA

iv. PROCESSING AND ANALYSIS OF ‘OMICS DATA

I. Technology
II. Sample Processing **
III. Transcriptomics Study Design
IV. Specification of Raw Data **
V. Data Normalization **
VI. Data filtering
VII. Identification and Removal of Low Quality or Outlying Datsets **

Content is platform-specific

Sections to be drafted by a **section workgroup** under guidance of a **section leader**

** *Emphasis on the use and description of Quality Control procedures / samples / performance metrics.*

Downstream Analysis Reporting Modules (DARMs)

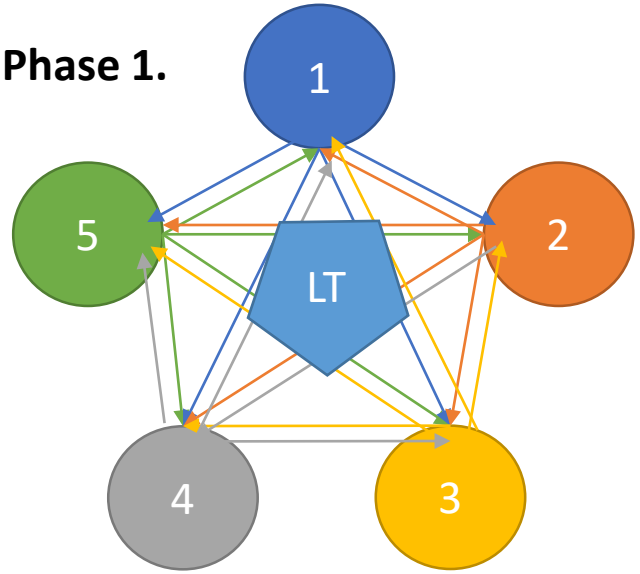
- **Downstream Analysis Reporting Modules (DARMs)**
  - Originally conceived by project leadership as a set of reporting templates complementary to the TRF.
  - Originally thought to be beyond the scope of current TRF project (i.e. follow-up work).
  - **DOES** prompt user to list all components of an analysis necessary to computationally reproduce results
  - **DOES NOT** tell the user which method, or iteration of a method, they should be using.
- **Identification of Differentially Expressed Genes (DEGs) will be piloted as the first DARM.**

Round Robin Case Study

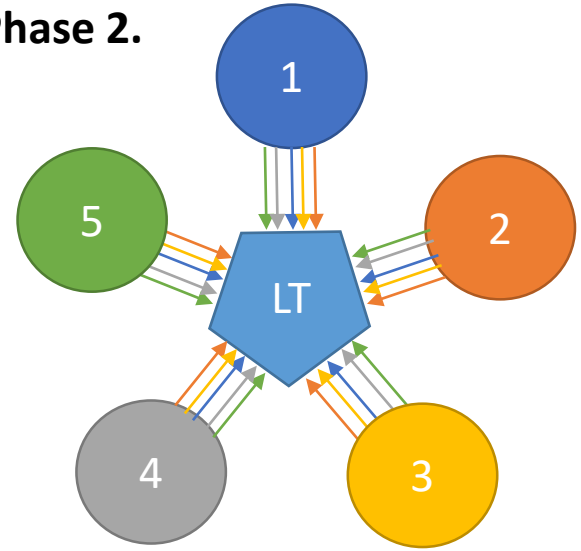
Objectives: Evaluate the utility of the TRF in fostering reproducibility of 'omics data analysis by different research groups.

Phase 1	**Step 1.**	Identify three (or more) analysis teams from various organizations.
	Step 2.	Coordinate with the leadership team to identify an existing dataset from each team
	Step 3.	Ask each team to: 1) Analyze their data & determine DEGs (no other instructions or restrictions). 2) Report DEGs and 3) Fill out the TRF describing what they did
Phase 2	**Step 4.**	Provide raw data and completed TRFs (blinded, sans DEG list) to other analysis teams
	Step 5.	Ask teams to: 1) Try and reproduce the analysis described in the TRF 2) Report DEGs to leadership team 3) Identify areas in the completed TRFs which were unclear
	Step 6.	Leadership team assesses concordance of DEG call results and report results back to analyses teams.
	Step 7.	Refine TRF (if necessary)

Phase 1.



Phase 2.



Section Workgroups

Each workgroup will consist of the following:

Title	Identity	Roles
Section Leads	Experiment — Raffaella Corvi [JRC] Microarray — Vikrant Vijay [NCTR] RNA-Seq — Florian Caiment [Maastricht] q-PCR array — Jason O'Brien [ECCC] TempO-Seq — Scott Auerbach [NTP] DARM.1 [DEG] — Lyle Burgoon [ERDC]	Coordinate workgroup activities Maintain draft of section Manage timelines for deliverables
Workgroup Members (n = 2-3)	See Next Slide	Contribute text and content for sections
"Floating" Facilitators	Joshua Harrill [USEPA] Carole Yauk [Health Canada]	Ensure consistency and cross-talk with other workgroups. Monitor progress in accordance with project timeline Foster discussion.
OECD Secretariat	Magda Sachana	Project administration / OECD liaison

All members of the TRF workgroup will have the opportunity to comment on each section.

Project group leads (Harrill & Yauk) will integrate sections into the final document.

Charge to Section Workgroups

iv. EXPERIMENT

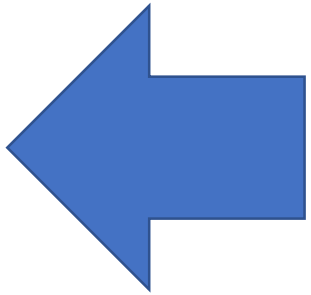
I. Study Rationale
II. Study Design
III. Subject / Test System Characteristics
IV. Test Article
V. Treatment Conditions
VI. Study Exit
VII. Sample Collection & Pre-processing
VIII. Sample Identification Codes
IX. Supporting Data Streams

iv. PROCESSING OF 'OMICS DATA [Microarray]

I. Technology
II. Sample Processing
III. Transcriptomics Study Design
IV. Specification of Raw Data
V. Data Normalization
VI. Data filtering
VII. Identification and Removal of Low Quality or Outlying Datasets

DARM.1 [DEGS]

I. ??????



<u>The section workgroups are tasked with:</u>

- Determining what information the user may list under each heading.
- Identify gaps (if any) that need to be added to the TRF structure.
- Determine the level of descriptive detail that is appropriate for each section
- "Beta test" the section using a couple of examples.
- Don't forget to use existing resources if available!

Project Timeline

Date	Milestone
April, 2018	Kickoff teleconference / recruiting for workgroups
May – June, 2018	Begin work on Introduction, Experiment, Microarray and DARM.1 modules
June, 2018	OECD WPHA & EAGMST Meeting – Project update (presentation)
Dec, 2018	First drafts of Introduction, Experiment and Microarray sections due OECD Winter Meeting
June, 2019	Near Final Draft of Introduction, Experiment and Microarray sections Kickoff of Round Robin Case Study for Microarray First drafts of RNA-Seq, PCR array, TempO-Seq due OECD Spring Meeting
Dec, 2019	Final document(s) – project completion OECD Winter Meeting

Progress To Date

Date	Milestone
February-April, 2018	Project leadership planning calls Drafting and circulation of TRF outline document
April, 2018	Kick-off teleconference with entire TRF working group • Solicitation of comment on TRF outline • Recruiting for section working groups • Addition of industry members
May, 2018	Second teleconference with entire TRF working group • Presentation of scoping statement • Follow-up on discussion points on document content / structure • Alignment of TRF with OECD Harmonized Templates (Alberto Martin, EFSA)
June, 2018	Kickoff TC for Experiment working group Kickoff TC for Microarray working group
July – Nov, 2018	Kickoff TC for DARM.1 Working Group Drafting of Experiment, Microarray and DARM.1 sections Monthly TC with each active working group.

Acknowledgements

- Leadership Team
  - Carole Yauk
  - Tim Gant
  - Magda Sachana

- OECD TRF Working Group Members

Questions?