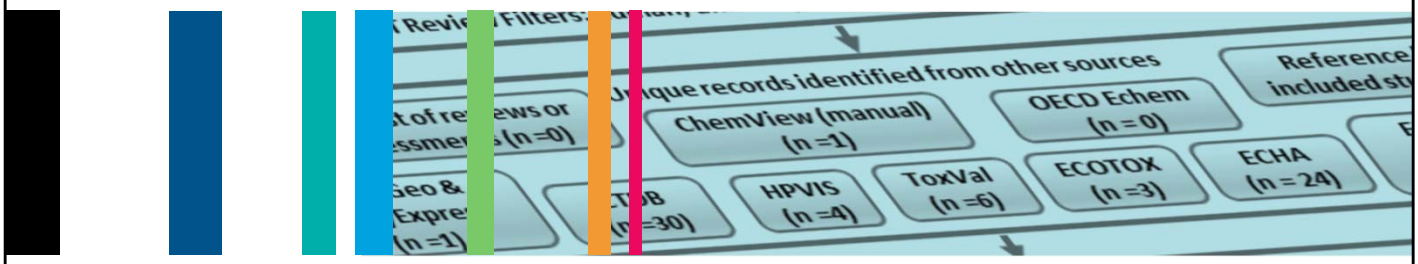
icf.com



we are ICF

Gray Literature Search

Nicole Vetter
Manager, Information Sciences

December 2, 2020

We will cover:

- **Introduction to gray literature**
  - Difference between gray literature and peer-reviewed literature
  - Reasons to use gray literature
  - Challenges associated with gray literature
  - Products (diagrams, tables, and templates) used to track and summarize results
- **Planning your gray literature search**
  - Commonly used sources of gray literature
  - Preparation includes considering project goals and organizing for flexibility and a consistent tracking method
- **Process and best-practices**
  - Tracking is critical to gray-literature information management
  - Applying the process to various sources

Introduction to gray literature

3

What is gray literature?

Gray literature refers to materials that are created "outside of traditional publishing and distribution channels." Examples include white papers, government documents, registration dossiers, evaluation reports, posters, presentations, conference proceedings, etc.
https://www.lib.sfu.ca/help/research-assistance/format-type/grey-literature

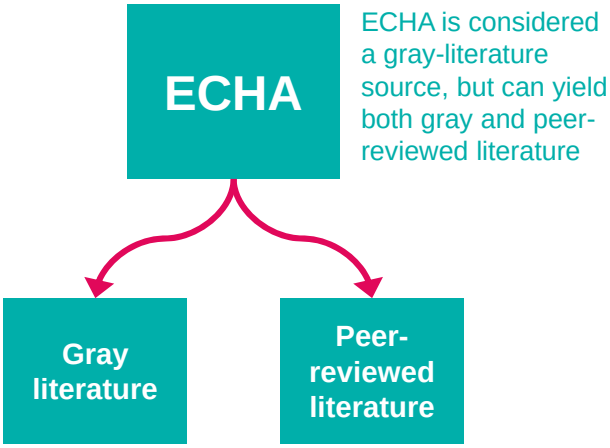
The Fourth International Conference on Gray Literature (GL '99) in Washington, DC, defined gray literature as "That which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers."
https://csulb.libguides.com/graylit

How does gray literature differ from peer-reviewed?

- Unlike peer-reviewed materials, gray literature is **not always assessed using quality measures**.
- Gray literature is **not consistently indexed** by bibliographic databases such as PubMed, Web of Science, etc.
- Citations may **not meet the same bibliographic standards** as commercially published documents.

Gray-literature sources vs. gray literature

- **Gray-literature sources:** **places where we expect to find gray literature**
  - These sources may still provide peer-reviewed literature
- **Gray literature:** **literature that has not been peer-reviewed for commercial publication**



ECHA is considered a gray-literature source, but can yield both gray and peer-reviewed literature

- Gray-literature sources, or places where we expect to find gray literature, can often also yield peer-reviewed literature.
- This is why tracking citation and study information is important throughout the gray-literature-search process; it helps facilitate de-duplication of any references that overlap with references found during the peer-reviewed search.
- Another example: A backwards search, or reference list review of included studies, is often included as a gray-literature source. The results of this search may yield gray and peer-reviewed literature. These results would also go through de-duplication for any references that overlap with those found during the peer-reviewed search.

Why use gray literature?

Goal: Comprehensive reference list which yields all available relevant data

- May need to supplement data from peer-reviewed literature with data from gray literature for data-poor chemicals
- Some chemicals may be well-studied by industry but not by academia
- Null or negative data can be found in gray literature
- Can provide a QC check against peer searches

Using gray literature can be like panning for gold:
It may be difficult to find, but there can be big payoffs when you find it!

ICF

7

- The goal for a given project is to create a master reference list that includes all possible unique references that yield relevant data or data of interest.
- Gray literature might be helpful for any of the reasons shown here.
- Gray literature might be hard to find but can be very useful when identified.

Challenges associated with gray literature

Incomplete Citations

Citations found in gray literature often consist of limited or incomplete information which makes it difficult to track and find sources.

Challenges detecting duplicates

No abstracts

No single trusted platform or repository

PubMed Citation

Per- and Polyfluoroalkyl Substance Exposure, Gestational Weight Gain, and Postpartum Weight Changes in Project Viva.

Mitro SD, Sagiv SK, Rifas-Shiman SL, Calafat AM, Fleisch AF, Jaacks LM, Williams PL, Oken E, James-Todd TM.

Obesity (Silver Spring). 2020 Oct;28(10):1984-1992. doi: 10.1002/oby.22933.

PMID: 32959518

Findings were similar after adjustment for all **PFAS**. Other **PFAS** were not associated with weight changes. Postpartum associations were stronger among women with higher prepregnancy BMI. ...CONCLUSIONS: Pregnancy **PFAS** were associated with greater gestational we ...

ECHA Citation

Reference	
Reference Type:	study report
Title:	Unnamed
Year:	1988
Report Date:	1988

8

- In general, gray literature often poses many unique challenges.
- With peer-reviewed literature, you typically have controlled bibliographic information that can be easy to use and reference. The incomplete nature of citations found in gray literature often makes it challenging to collect and track data.

Challenges associated with gray literature

Incomplete Citations

Challenges detecting duplicates

It can be difficult to compare references found within different sources due to limited information and inconsistent methods of citing references.

No abstracts

No single trusted platform or repository

ToxVal Export

toxval_type_original	toxval_numeric	toxval_units	sex	critical_effect	details_text
NOAEL	25	mg/kg-	male	fertility	ECHA IUCLID

ECHA Export

DataSource	Results	Results	Results	Results
Unnamed	male	NOAEL	25 mg/kg	fertility

ICF

9

- Another challenge that stems from having uncontrolled bibliographic information is being able to detect overlaps between references retrieved from different sources. Inconsistent citation methods and limited reference information can, therefore, make it difficult to distinguish between duplicate results.
- An additional form of duplication can be introduced by gray literature when a lab report or company submission to TSCA includes data that is also used in a peer-reviewed paper. Careful review of extracted data can help identify this kind of duplicated data.

Challenges associated with gray literature

Incomplete Citations

Challenges detecting duplicates

No abstracts

Unlike peer-reviewed literature, gray literature often does not have abstracts associated with the reference.

No single trusted platform or repository

PubMed provided citation

Title
Time and temperature effect on the digestive ripening of gold nanoparticles: Is there a cross over from digestive ripening to Ostwald ripening?

Author(s)
Sahu, P; Prasad, BL

PMID
25111614

DOI
10.1021/la500914j

Abstract
The effect of time and temperature on the gold nanoparticle sizes obtained by digestive ripening has been investigated. In digestive ripening a polydispersed colloid upon refluxing with a surface active ligand in a solvent, gets converted into a nearly monodisperse one. In this study a polydisperse gold nanoparticle system was heated in tertiary butyl toluene with hexadecanethiol at different temperatures viz. 60, 90, 120, 150 and 180 ºC for different time periods and the trends in particle size variations were recorded. At lower temperatures such as 60 ºC and 90 ºC, after the initial narrowing of size distribution, the particle sizes remain constant even though the refluxing step is continued for 24h substantiating the prevalence of digestive ripening process. However, at elevated temperatures (120, 150 and 180 ºC) particle sizes grow continuously indicating a deviation from the digestive ripening behaviour to an Ostwald ripening type phenomenon.

ECHA manually created citation

Title
Registration dossier: 1-naphthol (CASRN: 90-15-3): Developmental toxicity / teratogenicity: 002 Supporting | Experimental result

Author(s)
European Chemicals Agency :: ECHA

Year
2019

Language
English

URL
https://echa.europa.eu/registration-dossier/-/registered-dossier/22250/7/9/3/?documentUUID=5d18157d-860a-4f9f-b607-9738ddb1b39f
EXIT

10

- Abstracts, which are commonly available for peer-reviewed references, often do not exist with gray literature. This further makes it difficult to determine what types of data and information are present within a given reference.

Challenges associated with gray literature

Incomplete Citations

Challenges detecting duplicates

No abstracts

No single trusted platform or repository

There is no single repository that contains gray literature. Therefore, these references often need to be searched in multiple databases.

eChemPortal

ECHA EUROPEAN CHEMICALS AGENCY

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

EPA United States Environmental Protection Agency

ICF

11

- Since there is no single trusted platform or repository which houses these types of references, multiple databases are typically searched.
- There are other source-specific challenges that come with gray literature which will be discussed later on in the presentation.

Product examples

Other Sources Summary Table

- Shows complete information about the gray literature and other non-bibliographic database sources and the returned results

6.4.11. Summary table for other sources search results

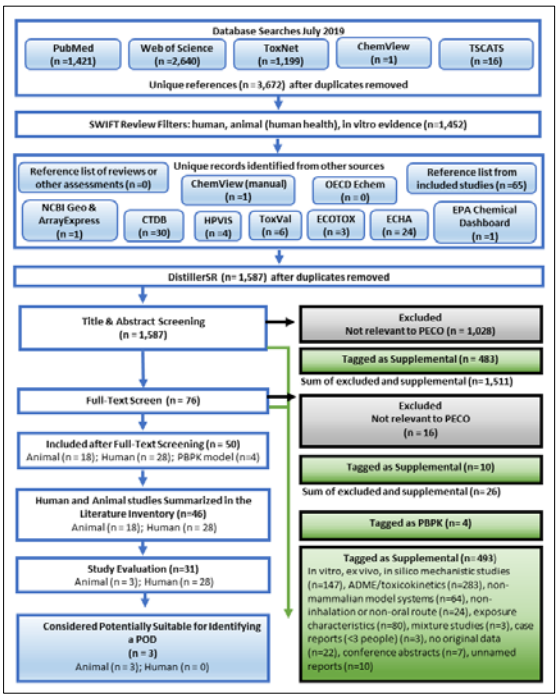
Source	Source Address	Search Terms	Search Date	Total Unique Number of Results Retrieved	Records not otherwise identified that were screened in DistillerSR
Review of reference lists studies considered relevant to PECO based on full-text screening.	NA	NA	NA	67	65
Review of reference lists from existing assessments (final or publicly available draft) or journal review articles that focused on human health	NA	NA	NA	3	0
EPA CompTox Chemicals Dashboard (ToxVal)	https://comptox.epa.gov/dashboard/dsstoxdb/results?abbreviation=TOXVAL_V5&search=DTXSID6021793#toxicity-values	90-15-3 (Results from Human Health: POD, Toxicity Value, Lethality Effect Level)	12/10/2019	21	6
ECHA	https://echa.europa.eu/information-on-chemicals/information-from-existing-substances-regulation	90-15-3	10/8/2019	53	24
EPA ChemView	https://chemview.epa.gov/chemview?tf=0&ch=90-15-3&su=2-5-6-7-37574985&as=3-10-9-8&ac=1-15-16-6378999&ma=4-11-1981377&tds=0&tdl=10&tas1=1&tas2=asc&tas3=undefined&tss=	90-15-3	9/19/2019	3	1
High Production Volume Challenge Database (HPVIS)	https://ofmpub.epa.gov/oppthpv/quicksearch.display?pChem=101850	90-15-3	9/19/2019	4	4

- How we make sense of the results we have collected.
- This example summary table includes unique counts of how many references were found vs how many were unique to the project.

Product examples

Study Flow Diagram

- Also called Literature Flow Diagram or PRISMA Flow Diagram
- Provides a visual narrative of the flow of information through the different steps of the Systematic Review process



13

- How we make sense of the results we have collected.
- The study flow diagram shows the number of references that were pulled into each step of the process and how they were categorized.

Planning your gray-literature search

Main sources of gray literature

Typically searched for various assessments (TSCA, CPAD, PPRTV, etc.)

- **ECHA**
- CompTox/ToxVal
- NTP CEBS
- ChemView
- OECD SIDS
- ECOTOX

ECHA
EUROPEAN CHEMICALS AGENCY

The **European Chemicals Agency (ECHA)** consists of registration dossiers used to identify data submitted by registrants. The amount of information provided for each substance varies and is obtained directly from companies' REACH registrations.

Registration dossiers contain valuable data on substances such as hazardous properties, safe uses, classifications, environmental fate, and ecotoxicological and toxicological information.

https://echa.europa.eu/information-on-chemicals/registered-substances

ICF

15

- We have listed sources we typically search depending on who is conducting the assessment (TSCA, CPAD, PPRTV, etc.). These slides include a brief description of the source, the organization that runs it, and an overview of information provided within each.
- The European Chemicals Agency (ECHA) consists of registration dossiers which includes data on substances such as hazardous properties, safe uses, classifications, environmental fate, and ecotoxicological and toxicological information.

Main sources of gray literature

Typically searched for various assessments (TSCA, CPAD, PPRTV, etc.)

- ECHA
- **CompTox/ToxVal**
- NTP CEBS
- ChemView
- OECD SIDS
- ECOTOX

EPA CompTox Chemicals Dashboard

The CompTox Chemicals Dashboard includes the Toxicity Values (ToxVal) database. ToxValDB collates publicly available toxicity dose–effect related summary values typically used in risk assessments.

These include:

- Point of Departure (POD) data
  collected from data sources within ACToR and ToxRefDB
- No-observed and lowest-observed (adverse) effect levels (NOEL/NOAEL, LOEL/LOAEL) data
  extracted from repeated dose toxicity studies submitted under REACH
- Reference dose and concentration values (RfDs and RfCs)
  from EPA's Integrated Risk Information System (IRIS)
- Dose descriptors
  from EPA's Provisional Peer-Reviewed Toxicity Values (PPRTV) documents

https://comptox.epa.gov/dashboard

ICF

16

- ToxValDB, which can be accessed through the CompTox Chemicals Dashboard, collates publicly available toxicity dose–effect related summary values typically used in risk assessments

Main sources of gray literature

Typically searched for various assessments (TSCA, CPAD, PPRTV, etc.)

- ECHA
- **CompTox/ToxVal**
- NTP CEBS
- ChemView
- OECD SIDS
- ECOTOX

EPA CompTox Chemicals Dashboard

The CompTox Chemicals Dashboard includes the Toxicity Values (ToxVal) database. ToxValDB collates publicly available toxicity dose–effect related summary values typically used in risk assessments.

These include:

- Acute toxicity information
  extracted from a number of different sources, including:
  - OECD eChemPortal,
  - ECHA (European Chemicals Agency),
  - NLM (National Library of Medicine) HSDB (Hazardous Substances Data Bank),
  - ChemIDplus via EPA TEST (Toxicity Estimation Software Tool)
  - EU JRC (Joint Research Centre) AcutoxBase.

ToxVal data can be accessed through the CompTox Chemicals dashboard or provided directly through CCTE.

https://comptox.epa.gov/dashboard

ICF

17

- The CompTox Chemicals Dashboard includes the Toxicity Values (ToxVal) database. ToxValDB collates publicly available toxicity dose–effect related summary values typically used in risk assessments.

Main sources of gray literature

Typically searched for various assessments (TSCA, CPAD, PPRTV, etc.)

- ECHA
- CompTox/ToxVal
- **NTP CEBS**
- ChemView
- OECD SIDS
- ECOTOX

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

The **Chemical Effects in Biological Systems (CEBS) database** houses data of interest to environmental health scientists. CEBS is a public resource and has received depositions of data from academic, industrial, and governmental laboratories. CEBS is designed to display data in the context of biology and study design, and permit data integration across studies for novel meta-analysis.

https://manticore.niehs.nih.gov/cebssearch/

ICF

18

- The Chemical Effects in Biological Systems (CEBS) database houses data of interest to environmental health scientists. CEBS is a public resource and has received depositions of data from academic, industrial, and governmental laboratories. CEBS is designed to display data in the context of biology and study design, and permit data integration across studies for novel meta-analysis.

Main sources of gray literature

Typically searched for various assessments (TSCA, CPAD, PPRTV, etc.)

- ECHA
- CompTox/ToxVal
- NTP CEBS
- **ChemView**
- OECD SIDS
- ECOTOX

EPA ChemView

ChemView is used to identify unpublished studies, information submitted to EPA under Toxic Substances Control Act (TSCA) Section 4 (chemical testing results), Section 8(d) (health and safety studies), Section 8(e) (substantial risk of injury to health or the environment notices), and FYI (voluntary documents).

Other databases accessible via ChemView include EPA's High Production Volume (HPV) Challenge database, and the Toxic Release Inventory (TRI) database.

https://chemview.epa.gov/chemview

ICF

19

- ChemView is used to identify unpublished studies and information submitted to EPA under Toxic Substances Control Act (TSCA). This also includes access to information from HPV & TRI databases.

Main sources of gray literature

Typically searched for various assessments (TSCA, CPAD, PPRTV, etc.)

- ECHA
- CompTox/ToxVal
- NTP CEBS
- ChemView
- **OECD SIDS**
- ECOTOX

OECD
BETTER POLICIES FOR BETTER LIVES

The **OECD Screening Information Data Sets (SIDS) project** is a cooperative effort of OECD countries designed to collect information on High Production Volume chemicals. One of the goals of the SIDS project is to locate the data needed for the initial assessment of these chemicals and to generate data which were lacking.

This source is especially useful for capturing government assessments and houses public health and environmental data that can be found across multiple government databases.

https://hpvchemicals.oecd.org/ui/Default.aspx

ICF

20

- The OECD Screening Information Data Sets (SIDS) project is a cooperative effort of OECD countries designed to collect information on High Production Volume chemicals. One of the goals of the SIDS project is to locate the data needed for the initial assessment of these chemicals and to generate data which were lacking.
- OECD is also a great one stop for government assessments.

Main sources of gray literature

Typically searched for various assessments (TSCA, CPAD, PPRTV, etc.)

- ECHA
- CompTox/ToxVal
- NTP CEBS
- ChemView
- OECD SIDS
- **ECOTOX**

EPA ECOTOX Knowledgebase

The **ECOTOXicology Knowledgebase (ECOTOX)** is a comprehensive, publicly available knowledgebase providing single chemical environmental toxicity data on aquatic life, terrestrial plants and wildlife.

https://cfpub.epa.gov/ecotox/index.cfm

ICF

21

- The ECOTOXicology Knowledgebase (ECOTOX) is a comprehensive, publicly available knowledgebase providing single chemical environmental toxicity data on aquatic life, terrestrial plants and wildlife.

Secondary sources of gray literature

Typically incorporated for data-poor chemicals

- **CTDB**
- **GEO**

The Comparative Toxicogenomics database (CTDB) includes references that report chemical-gene interactions and information that can enrich the inventory of mechanistic references.

http://ctdbase.org/

22

- **The following consists of sources we typically search for data-poor chemicals.**
- The Comparative Toxicogenomics database (CTDB) includes references that report chemical-gene interactions and information that can enrich the inventory of mechanistic references.

Secondary sources of gray literature

Typically incorporated for data-poor chemicals

- CTDB
- GEO

The Gene Expression Omnibus includes public data from gene expression studies which can be used to generate mechanistic evidence, predict adverse outcomes, and derive toxicity values depending on the design of gene expression studies which produced these data.

https://www.ncbi.nlm.nih.gov/geo/

- The Gene Expression Omnibus includes public data from gene expression studies which can be used to generate mechanistic evidence, predict adverse outcomes, and derive toxicity values depending on the design of gene expression studies which produced these data.

Search preparation considerations

Project goals

- The gray-literature search process may differ between projects depending on project goals and information of interest
- Project goals and data of interest may influence:
  - How many and which sources are searched for gray literature
  - How gray literature from each source is recorded and tracked
  - Which data are extracted from each gray-literature reference

24

- There are many things you will want to consider before beginning a gray-literature search. Arguably the most important of these are your project goals, which will likely influence which sources you search and how information is recorded or extracted from those sources.
    - Depending on the sources used in the search process, additional fields may be required to help with the deduplication process.
    - It is important to locate and extract data that is relevant to your project and scope. Some fields may not be necessary to include from certain sources depending on project-specific interests.

Search preparation considerations

Project goals

- The gray-literature search process **may differ between projects** depending on project goals and information of interest
- Other project considerations:
  - Contingent on budget/time restraints
  - Number of chemicals included in the search
  - Type of references (company vs government)

25

- The types and number of sources searched could be affected by budget and time restraints, as well as the amount chemicals that are included in the search.
- It is also important to consider the amount of data that typically exists for a given chemical. Some chemicals may not have much literature associated with them whereas others are data-rich.
- With limited time, you may want to search ToxVal which already collates data from other sources; if you have a data-poor chemical, you may want to search additional sources to find as much data as you can.

Search preparation considerations

Project goals Example

A project interested in extracting **hazard-identification data** might search these sources:

- ECHA – tox info
- CompTox/ToxVal

And record or extract data from each reference on **study design and results** (e.g., PODs) in order to help identify duplicate references

A project interested in extracting **exposure data** might search these sources:

- ECHA – exposure info
- USGS Monitoring Data

And record or extract data from each reference on **media and results** (e.g., exposure level) to help identify duplicate references

26

- For example, a project focusing on hazard-identification data might search ECHA and be interested in the **toxicity information**, and also search for results in ToxVal.
    - Fields of interest for this project would also likely include study design and results information (like PODs, endpoints, species, etc.) from the gray literature to help identify duplicate references later in the process.
- A project interested in exposure data might look different; they would likely want to search ECHA as well but would be interested in the **exposure information**. They would also likely search something like the USGS Monitoring database.
    - The fields of interest in that project would focus on different information from the identified gray literature and would use those different data to find duplicates later in the process.

Track!

Search preparation considerations

Tracking

- **It is critical to maintain results collected from each source in a centralized location that can be cross-checked against a master reference list.**
  - Gray-literature tracking might overlap with data extraction because study details might be necessary to help identify duplicates
  - ICF has developed templates (which can be modified as necessary) to help retrieve, deduplicate, and prepare references for review.

ICF

27

- We included a small ‘tracking’ icon on slides of the presentation that focus on the tracking process and best practices since it’s so important.
- Carefully tracking the sources searched and the citations and data extracted from those sources is critical to successful integration of gray literature into a systematic review process.
- ICF has developed some templates that could be used as starting points and edited for future use.

Track!

Search preparation considerations

Tracking Example

Excel tracking files are often composed of large tables and many worksheets if multiple gray-literature sources are searched

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Search_ID	Search_Type	Completed_Date	ChemShortname	DSSTox_Substance_ID	CAS_Number	PREFERRED_NAME	Chemical Searched	Has Results?	Result ID Number	Url Searched	Search Terms	Result URL	Filename
11	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-5	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-5.pdf	
12	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-5	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-5.pdf	
13	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-6	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-6.pdf	
14	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-7	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-7.pdf	
15	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-8.pdf	
16	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-8.pdf	
17	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-8.pdf	
18	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-8.pdf	
19	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-8.pdf	
20	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-2-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-2-8.pdf	
21	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3.pdf	
22	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-1	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-1.pdf	
23	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-2	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-2.pdf	
24	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-3	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-3.pdf	
25	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-4	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-4.pdf	
26	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-4	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-4.pdf	
27	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-4	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-4.pdf	
28	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-5	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-5.pdf	
29	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-6	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-6.pdf	
30	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-7	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-7.pdf	
31	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-8.pdf	
32	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-8.pdf	
33	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-8.pdf	
34	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-8	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-8.pdf	
35	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-9	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-9.pdf	
36	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-9	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-9.pdf	
37	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-9	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-9.pdf	
38	3057	Manual	12/18/2019	PFAS_0001	DTXSID0022134	920-66-1	2H-Perfluoro-2-propanol	1,1,1,3,3,3-hexafluoro	Yes	3057-PFAS_0001-3-9	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0001-3-9.pdf	
39	3057	Manual	12/3/2019	PFAS_0012	DTXSID3052620	86479-06-3	Hexaflumuron	1-(3,5-dichloro-4-(1,1,	Yes	3057-PFAS_0012-1	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0012-1.pdf	
40	3057	Manual	12/3/2019	PFAS_0012	DTXSID3052620	86479-06-3	Hexaflumuron	1-(3,5-dichloro-4-(1,1,	Yes	3057-PFAS_0012-1-1	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0012-1-1.pdf	
41	3057	Manual	12/3/2019	PFAS_0012	DTXSID3052620	86479-06-3	Hexaflumuron	1-(3,5-dichloro-4-(1,1,	Yes	3057-PFAS_0012-1-2-1	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0012-1-2-1.pd	
42	3057	Manual	12/3/2019	PFAS_0012	DTXSID3052620	86479-06-3	Hexaflumuron	1-(3,5-dichloro-4-(1,1,	Yes	3057-PFAS_0012-1-2-2	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0012-1-2-2.pd	
43	3057	Manual	12/3/2019	PFAS_0012	DTXSID3052620	86479-06-3	Hexaflumuron	1-(3,5-dichloro-4-(1,1,	Yes	3057-PFAS_0012-1-3	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0012-1-3.pdf	
44	3057	Manual	12/3/2019	PFAS_0012	DTXSID3052620	86479-06-3	Hexaflumuron	1-(3,5-dichloro-4-(1,1,	Yes	3057-PFAS_0012-1-4	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0012-1-4.pdf	
45	3057	Manual	12/3/2019	PFAS_0012	DTXSID3052620	86479-06-3	Hexaflumuron	1-(3,5-dichloro-4-(1,1,	Yes	3057-PFAS_0012-1-5	https://echa.eu	[CAS number]	https://echa.eu/3057-PFAS_0012-1-5.pdf	

ICF

28

- Gray-literature tracking files are often large – the Excel files usually contain many tables and worksheets.
- These include tabs for each source and results, working tabs for reviewing information or converting reported data to aid deduplication, and a master tab that brings it all together.

Track!

Search preparation considerations

Search.ID	Search_Type	Completed_Date	ChemShortname	DSSTox_Substance_ID	CAS Number	Chemical.Searched	Has Results?	Result.ID.Number	Url.Searched	Search.Terms	Result.URL	Filename

Begin with a blank file with columns you anticipate needing

Search.ID	Search_Type	Completed_Date	ChemShortname	DSSTox_Substance_ID	CAS Number	Chemical.Searched	Has Results?	Result.ID.Number	Url.Searched	Search.Terms	Result.URL	Filename
3057	Manual								https://echa.eu	[CAS number]		
3057	Manual								https://echa.eu	[CAS number]		
3057	Manual								https://echa.eu	[CAS number]		
3057	Manual								https://echa.eu	[CAS number]		
3057	Manual								https://echa.eu	[CAS number]		
3057	Manual								https://echa.eu	[CAS number]		
3057	Manual								https://echa.eu	[CAS number]		

Fill in available information after performing preliminary search

Search.ID	Search_Type	Completed_Date	ChemShortname	DSSTox_Substance_ID	CAS Number	Chemical.Searched	Has Results?	Result.ID.Number	Url.Searched	Search.Terms	Result.URL	Filename
3057	Manual	12/6/2019	PFAS_0079	DTXSID5060982	677-69-0	1,1,1,2,3,3,3-heptafluoro	Yes	3057-PFAS_0079-1	https://echa.eu	[CAS number]	https://echa.eu	3057-PFAS_0079-1.pdf
3057	Manual	12/6/2019	PFAS_0079	DTXSID5060982	677-69-0	1,1,1,2,3,3,3-heptafluoro	Yes	3057-PFAS_0079-1-1	https://echa.eu	[CAS number]	https://echa.eu	3057-PFAS_0079-1-1.pdf
3057	Manual	12/6/2019	PFAS_0130	DTXSID10171538	1841-46-9		No		https://echa.eu	[CAS number]		
3057	Manual	12/18/2019	PFAS_0001	DTXSID1022134	920-66-1	1,1,1,3,3,3-hexafluoropro	Yes	3057-PFAS_0001-2-3-1	https://echa.eu	[CAS number]	https://echa.eu	3057-PFAS_0001-2-3-1.pdf
3057	Manual	12/18/2019	PFAS_0001	DTXSID1022134	920-66-1	1,1,1,3,3,3-hexafluoropro	Yes	3057-PFAS_0001-2-3-2	https://echa.eu	[CAS number]	https://echa.eu	3057-PFAS_0001-2-3-2.pdf
3057	Manual	12/18/2019	PFAS_0001	DTXSID1022134	920-66-1	1,1,1,3,3,3-hexafluoropro	Yes	3057-PFAS_0001-2-4-1	https://echa.eu	[CAS number]	https://echa.eu	3057-PFAS_0001-2-4-1.pdf
3057	Manual	12/18/2019	PFAS_0001	DTXSID1022134	920-66-1	1,1,1,3,3,3-hexafluoropro	Yes	3057-PFAS_0001-2-4-2	https://echa.eu	[CAS number]	https://echa.eu	3057-PFAS_0001-2-4-2.pdf

Complete extraction

ICF

29

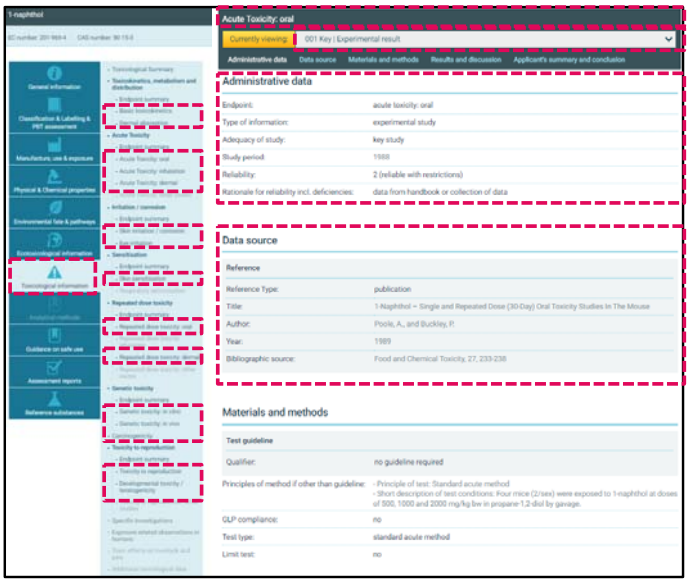
- This is only a snapshot of a small portion of the tracking sheet meant to serve as an example of what the final product may look like.
- You will fill in the columns as you retrieve information.
- As you begin identifying references, you may need to add columns for additional information of interest found in a particular source.

Track!

Search preparation considerations

Tracking Example

Defining data of interest to extract from gray-literature sources can help determine tracking needs



30

- Going back to the importance of determining project goals and data of interest: making these determinations up front can aid with determining the best way to organize a tracking file/system.

Search preparation considerations

Flexibility

- **The gray-literature search process should be considered iterative and should evolve as new information and scenarios that require improvements are encountered.**

Support

- **ICF can provide additional support, as needed, at any step of the process.**

- It is important to note that this is a constantly evolving process that has, and will continue to be, improved over time.
- ICF is always willing to help at any stage of the process and can provide support as needed.

Gray-literature search process and best-practices

32

- Now we will cover the gray-literature search process and best-practices.

Search process and tracking

Use this process to guide the search and track results:

Step 0: Consider project goals to determine plan-of-action and prepare tracking file (see previous section)

Step 1: Identify chemical search terms used for each source

Step 2: Perform initial search to determine what data exists in each source

Step 3: Capture citation information and relevant data

Step 4: De-duplicate and add new references to project database

Search process and tracking

Use this process to guide the search and track results:

Step 0: Consider project goals to determine plan-of-action and prepare tracking file (see previous section)

Step 1: Identify chemical search terms used for each source

Step 2: Perform initial search to determine what data exists in each source

Step 3: Capture citation information and relevant data

Step 4: De-duplicate and add new references to project database

Prepare tracking file

Step 0

Track!

Create a tab in the tracking file for each source

Incorporate sections that will be extracted in the tracking file:

- **General** fields consisting of search information, chemical information, and result identifiers
- **Source specific** fields consisting of data of interest
- **Reference** fields used for tracking and deduplication efforts

General Fields					Source Specific Fields			Reference Tracking			
Search Date	Chemical Name	Search Term	URL	Filename	General Info	Admin Data	Data Source	Citation Provided?	Citation Information	Found in Peer?	HERO ID

35

- These fields are generally incorporated in the tracking sheet for each source and contain information that is consistent across each chemical searched.
- We will add columns and sheets to the tracking file as the search progresses.

Prepare tracking file

Step 0

Track!

Create a tab in the tracking file for each source

Repeat the process for every source using the types of information available within each database.

Tabs	
Master Chemical List	A master list of all chemicals with their identifiers and synonyms.
1Nap ECHA Results	Output for the 1-Naphthol search in ECHA.
1Nap ToxVal References	A comprehensive list of references for the 1-naphthol ToxVal outpu
1Nap CEBS References	A list of references for 1-Naphthol pulled from CEBS.

ReadMe | 1NAP ECHA Results | 1NAP ToxVal References | 1NAP CEBS References

ReadMe tabs are also useful to incorporate and can provide more detailed information about each tab

Tabs	
Master Chemical List	A master list of all chemicals with their identifiers and synonyms.
1Nap ECHA Results	Output for the 1-Naphthol search in ECHA.
1Nap ToxVal References	A comprehensive list of references for the 1-naphthol ToxVal outpu
1Nap CEBS References	A list of references for 1-Naphthol pulled from CEBS.

ReadMe | 1NAP ECHA Results | 1NAP ToxVal References | 1NAP CEBS References

Search process and tracking

Use this process to guide the search and track results:

Step 0: Consider project goals to determine plan-of-action and prepare tracking file (see previous section)

Step 1: Identify chemical search terms used for each source

Step 2: Perform initial search to determine what data exists in each source

Step 3: Capture citation information and relevant data

Step 4: De-duplicate and add new references to project database

Identify chemical search terms used for each source

Step 1

Identify chemical terms

For chemical(s) of interest, gather information on:

- Chemical name
- Synonyms
- CASRN
- DTXSID of interest

Databases used to retrieve identifiers

- CompTox Dashboard
- ChemID*plus*

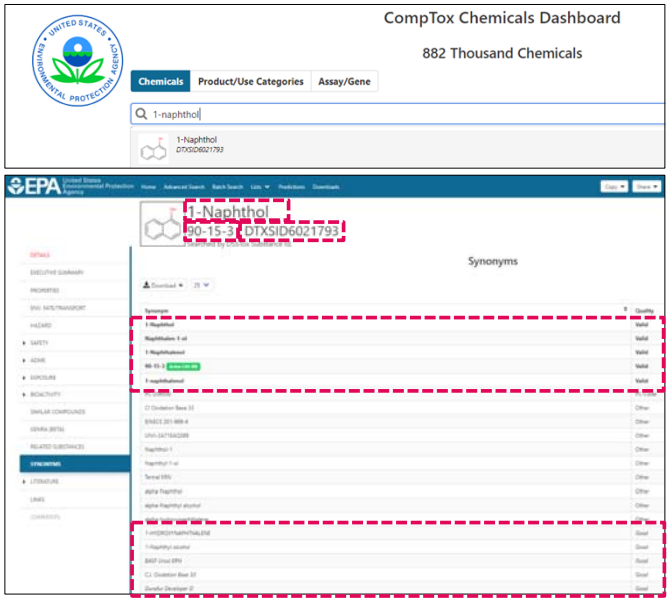
38

- It is important to gather chemical search terms up front before beginning the search of gray-literature sources.
- ICF typically searches CompTox Dashboard and ChemID*plus* to locate and retrieve chemical identifiers.

Identify chemical search terms used for each source

Step 1

Identify chemical terms using CompTox Dashboard



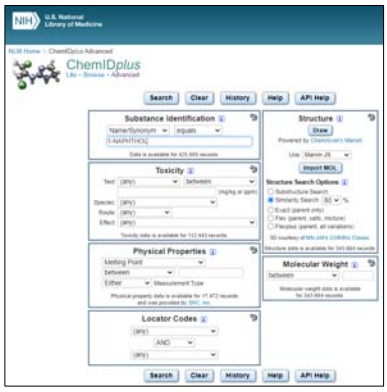
39

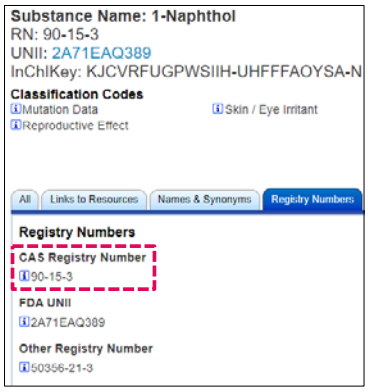
- The CompTox Dashboard can be searched to obtain the CAS number, DTXSID, and synonyms for each chemical.
- Although the CAS and DTXSID numbers are typically used in the chemical searches, the synonyms are more frequently used during the peer-reviewed literature search and screening process. Therefore, it's good practice to keep a running list of these synonyms.
- Depending on your search, you may want to limit the synonyms obtained through the CompTox Dashboard to those categorized as "valid", "good", and "fair".

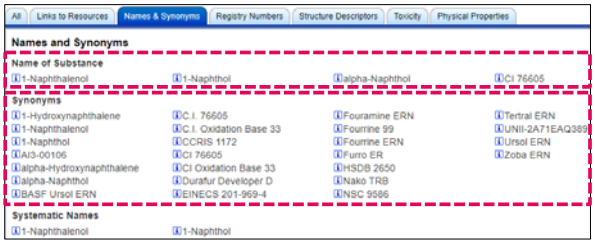
Identify chemical search terms used for each source

Step 1

Identify chemical terms using ChemID*plus*







40

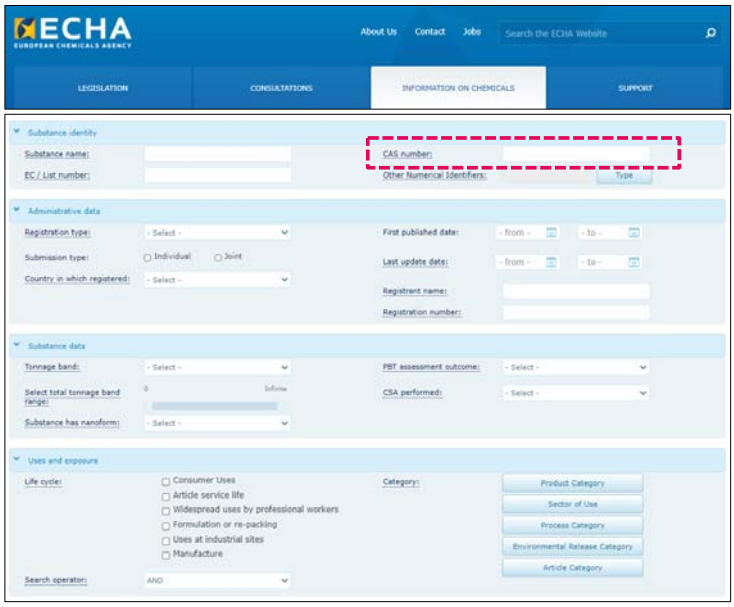
- The ChemID*plus* database can be searched to obtain information on both the chemical CAS number and synonyms.
- In the examples provided on this slide and the previous slide for 1-Naphthol, the synonyms retrieved from both ChemID*plus* and CompTox will be compiled.

ECHA example

Identify chemical search terms used for each source

Step 1

Identify the search terms used in each source



41

- Some sources may be limited to searching using the CASRN, DTXSID, chemical name, or synonym.
- The example here shows an ECHA chemical search by CASRN.
- ICF typically uses the CAS number field to retrieve results in ECHA, but other identifiers can also be used to search the database.

Identify chemical search terms used for each source

Step 1

Notes and best practices

- **CASRNs can be misleading** or incorrect, or can exist in multiple for the same chemical; DTXSIDS are becoming more frequently used
  - Some sources may **restrict** the types of identifiers you can use to search the database (e.g., ECHA does not allow searching by DTXSID)
- CompTox Dashboard and ChemID*plus* may provide **different synonym lists**; synonyms from both sources should be compiled
- Depending on the search, it may be necessary to **limit the synonyms obtained through the CompTox Dashboard** to those categorized as “valid”, “good”, and “fair”
- Using **correct chemical information** to conduct a search is the key to success

42

- Restricted identifiers: e.g., ECHA does not allow search by DTSXID

Identify chemical search terms used for each source

Step 1

Track!

Tracking

Incorporate chemical-information columns in tracking sheet

Chemical-specific identifiers

ChemShortname	DSSTox_Substance_iD	CAS Number	Chemical.Searched	Result.ID.Number	Search.Terms	Result.URL	Filename
1NAP	DTXSID6021793	90-15-3	1-naphthol		[CAS number]		

Abbreviated or short name for chemical

43

- Once chemical information has been identified, you can incorporate columns of chemical-specific identifiers and naming conventions used to track results into the tracking sheet.

Search process and tracking

Use this process to guide the search and track results:

Step 0: Consider project goals to determine plan-of-action and prepare tracking file (see previous section)

Step 1: Identify chemical search terms used for each source

Step 2: Perform initial search to determine what data exists in each source

Step 3: Capture citation information and relevant data

Step 4: De-duplicate and add new references to project database

Perform initial search to determine what data exists in each source

Step 2

Determine available data for each source

An **initial search** of each source should help determine:

- Whether data are available
- What type(s) of data are available
- How to capture files such as documents, reports, and registration dossiers
- How available data should be tracked

It is important to **become familiar with the way each source provides data** and results.

45

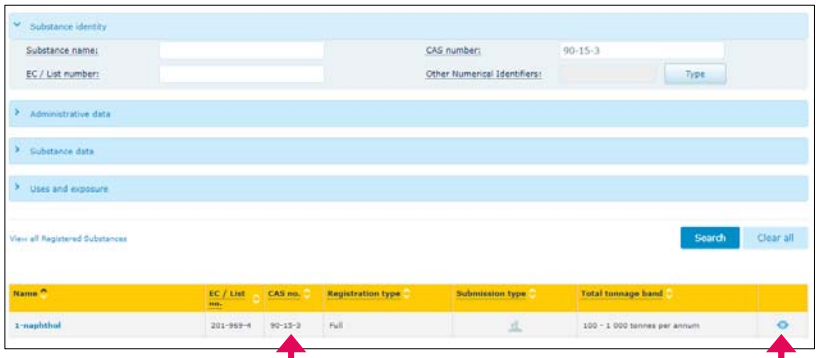
- Before you start downloading results, confirm which pieces of data are available from your sources, plan how you will capture the information (export or create PDF of web pages), and think about how the data will be tracked.

ECHA example

Perform initial search to determine what data exists in each source

Step 2

Determine available data in ECHA



Verify the CASRN under the CAS no. matches chemical of interest

Click on the "eye" icon to access the dossier

46

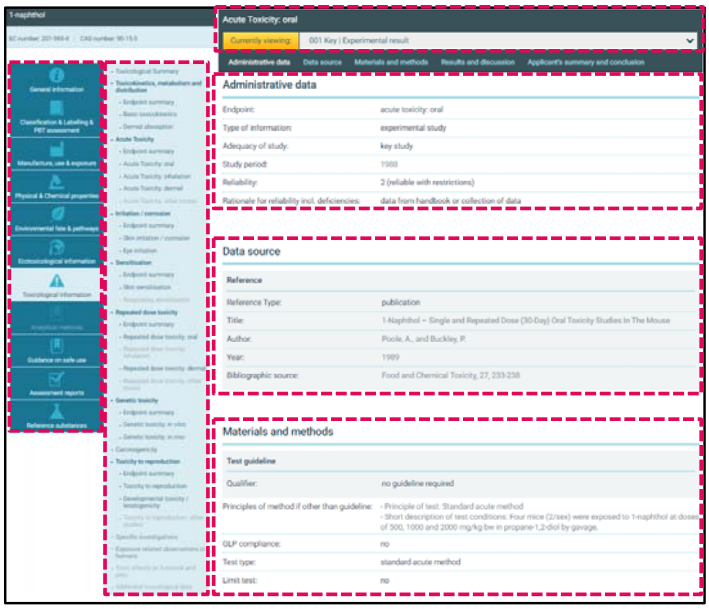
- This example shows search results in ECHA after entering a CASRN of interest; results populate below the search fields.
- Verify that the CASRN matches the chemical of interest. Sometimes different chemicals classified under the same chemical group can appear in the results.
- Some substances may also have multiple results for different submission and registration types. It is important to capture information from all dossiers of the chemical of interest.

ECHA example

Perform initial search to determine what data exists in each source

Step 2

Determine available data in ECHA



47

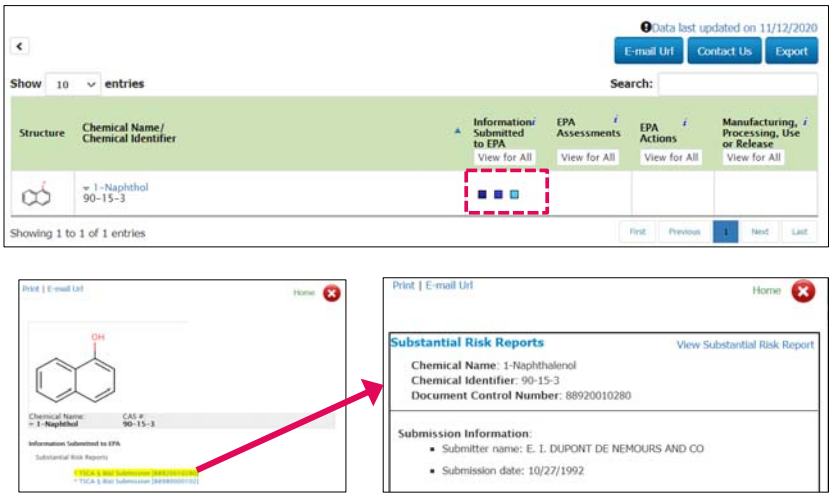
- Once you have opened the dossier, you will see the types of information available on the left panel. This example expands on the Toxicological Information tab, but other information is also available.
- Review the available endpoints and begin opening each nested link to see what data is available.

ChemView example

Determine available data in ChemView

Perform initial search to determine what data exists in each source

Step 2

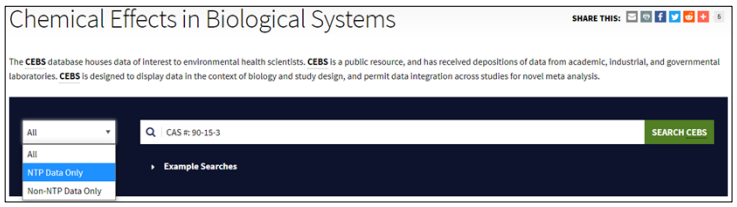


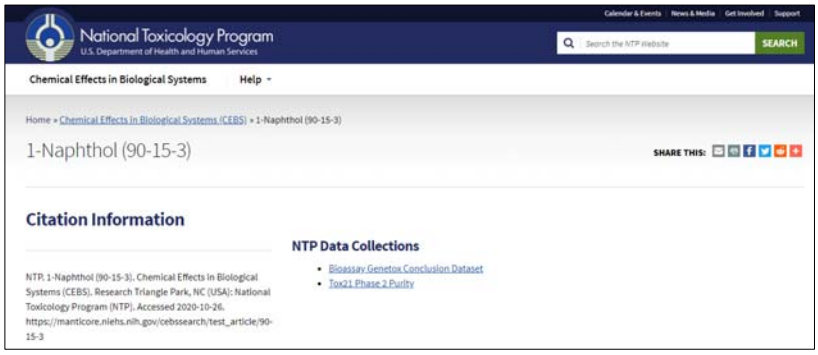
48

- This example shows search results in ChemView.
- There are three types of submissions available for this chemical (linked by the boxes): Substantial Risk Reports, Health and Safety Studies, and High Production Volume Information Systems.
- The types of information submitted to EPA can differ between chemicals depending on what is available. The images below the initial search entry show one type of submission - Substantial Risk Reports.

CEBS example

Determine available data in CEBS





Perform initial search to determine what data exists in each source

Step 2

49

- This example shows search results in CEBS, which include document citations as well as data collections.
- CEBS requires **specific search syntax** to retrieve results for a given chemical.
- We apply the NTP Data Only filter when conducting searches in this database.

Perform initial search to determine what data exists in each source

Step 2

Notes and best practices

- It is critical to **record 'no results'** from a given source for tracking purposes
- It is important to note **how each source provides data** and results of interest
- If possible, **assess any overlap** in results returned from multiple sources and **'triage' accordingly**
  - For example, ToxVal strives to aggregate data from multiple sources, so this may be a good starting place for a gray-literature search to streamline the de-duplication process after results have been identified from multiple sources
- ECHA-specific
  - Verify that the **CASRN matches the chemical of interest**; sometimes data from different chemicals can appear in the results
- CEBS-specific
  - CEBS requires a **specific search syntax** to retrieve results for a given chemical

50

- Record the number of results for each source, including 'no results' for chemicals that do not have data available.
- Identifying how each source provides data informs how to set up the tracking sheet.
- Triage the order of sources searched to assist in de-duplication.
- Source-specific examples:
  - ECHA – need to confirm the CASRN
  - CEBS – requires specific syntax

Perform initial search to determine what data exists in each source

Step 2

Track!

Tracking

- Data available in the **initial search should inform tracking** organization
- In an Excel file, typically:
  - Each source should have its own worksheet
  - Each field of interest from each source should occupy a column of a table
  - Each record would occupy a row of a table
- Hazard-specific:
  - Each data-source and point-of-departure-value combination should be extracted as a separate row

Search process and tracking

Use this process to guide the search and track results:

Step 0: Consider project goals to determine plan-of-action and prepare tracking file (see previous section)

Step 1: Identify chemical search terms used for each source

Step 2: Perform initial search to determine what data exists in each source

Step 3: Capture citation information and relevant data

Step 4: De-duplicate and add new references to project database

Capture citation information and relevant data

Step 3

Track!

Record relevant information

- This step **mostly involves tracking**; it is critical to stay organized and be comprehensive
- **Capture relevant record information and data** in the prepared tracking file once setup is complete (previous step)

53

- The initial searches performed in the previous step inform how to set up the tracking.

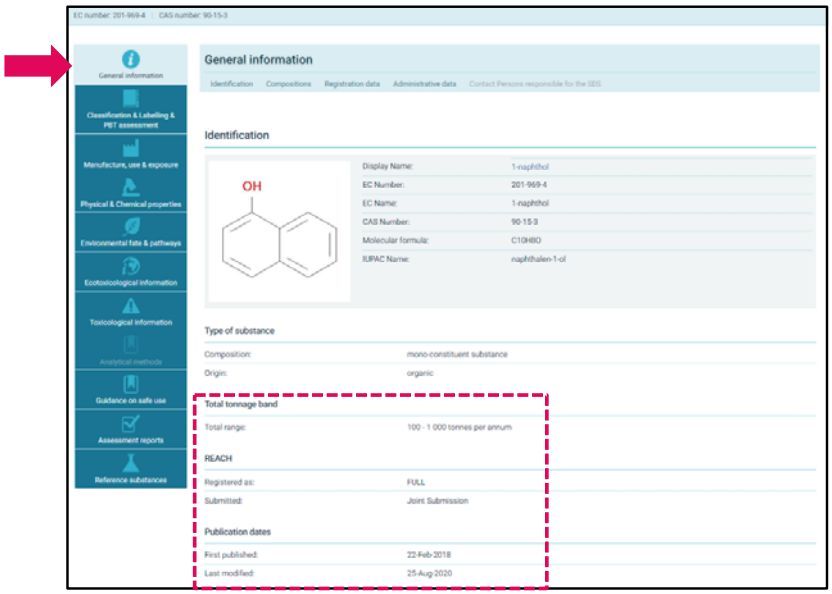
ECHA example

Capture citation information and relevant data

Step 3

Track!

Record relevant information from ECHA



54

- Information from the General Information tab of an ECHA dossier is always extracted as it contains valuable data on the submission and type of data published.
- Since these dossiers often change and are considered to contain “live” data, it is important to capture the last modified date to track the exact date information was accessed.
- Data extracted from the General Information tab will be identical for all rows of the tracking file within a **specific chemical dossier submission**.

ECHA example

Capture citation information and relevant data

Step 3

Track!

Record relevant information from ECHA

Total tonnage band	
Total range:	100 - 1 000 tonnes per annum
REACH	
Registered as:	FULL
Submitted:	Joint Submission
Publication dates	
First published:	22-Feb-2018
Last modified:	25-Aug-2020

TotalTonnageBand_TotalRange	100 - 1 000 tonnes per annum
Reach_Registration	FULL
Reach_Submission	Joint Submission
Publication_FirstSubmitted	22-Feb-18
Publication_LastUpdate	25-Aug-20

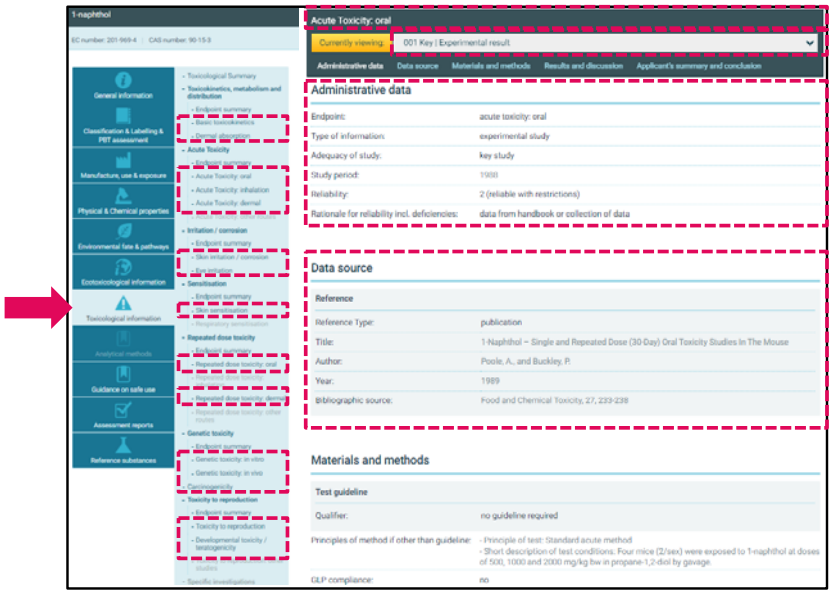
55

- Once all fields of interest are identified and corresponding columns added to the tracking sheet, you can begin extracting information from each section in the dossier.

ECHA example

Track!

Record relevant information from ECHA



Capture citation information and relevant data

Step 3

- For human health data, the Toxicological Information tab typically contains all the information we are interested in extracting.
- Each nested link under this tab would get its own row in the tracking sheet (e.g., acute tox: oral, acute tox: inhalation).
    - We are not interested in summary pages since the information within these will be captured under the other endpoints mentioned for each category.
- Although the details you choose to extract may be unique based on the project, ICF typically pulls data in the fields outlined in this example.
    - Depending on the endpoint, there may be additional sections that are not captured in this slide.
    - It is important to explore different endpoints to see what types of information are provided under each.
- Some endpoints may not have data for each column of the tracking sheet based on the information that is available.

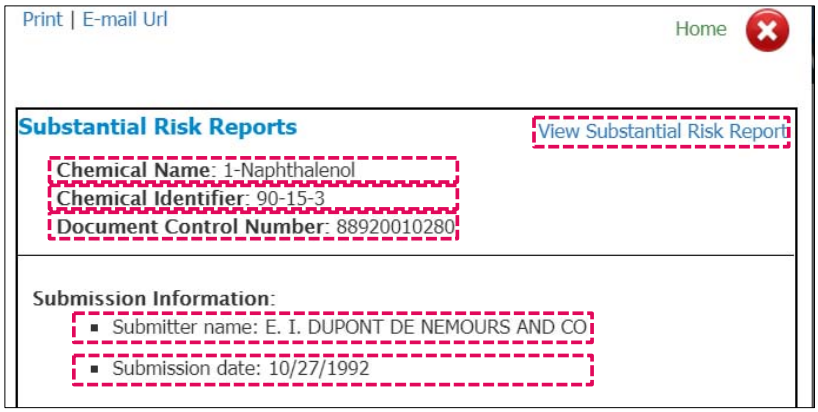
ChemView example

Capture citation information and relevant data

Step 3

Track!

Record relevant information from ChemView



57

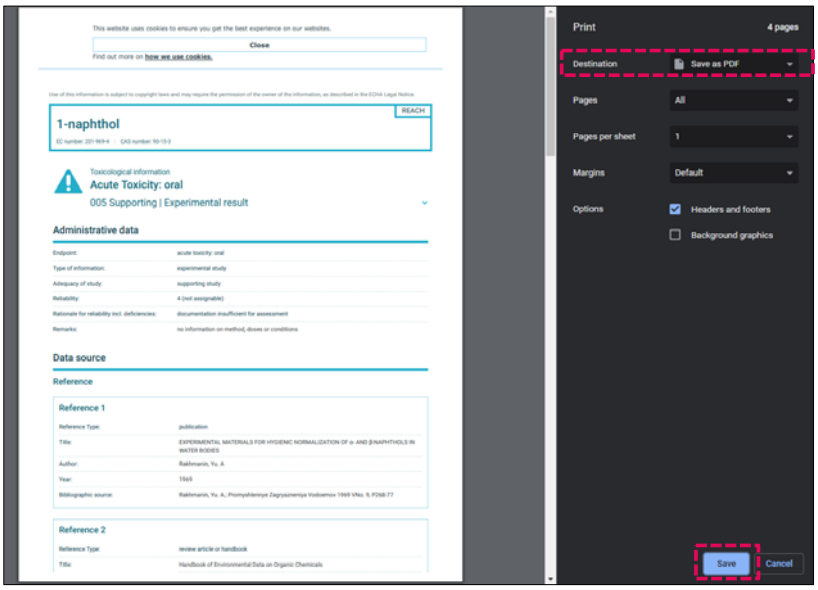
- As before, capture fields of interest and incorporate them in the tracking sheet for further extraction.
- Click on the “View Substantial Risk Report” link to access the downloadable report.

ECHA example

Capture citation information and relevant data

Step 3

‘Print’ pages from data sources



58

- Once extraction is complete, save each entry page as an individual PDF.
- Citations will be created based on the information within these PDFs, and HERO IDs will be assigned to these references.
- This is particularly important for any source that is considered ‘live’ data or may be updated regularly (such as ECHA).
- You can click on the “print” icon at the top of the page and save each file using an organized file-naming convention

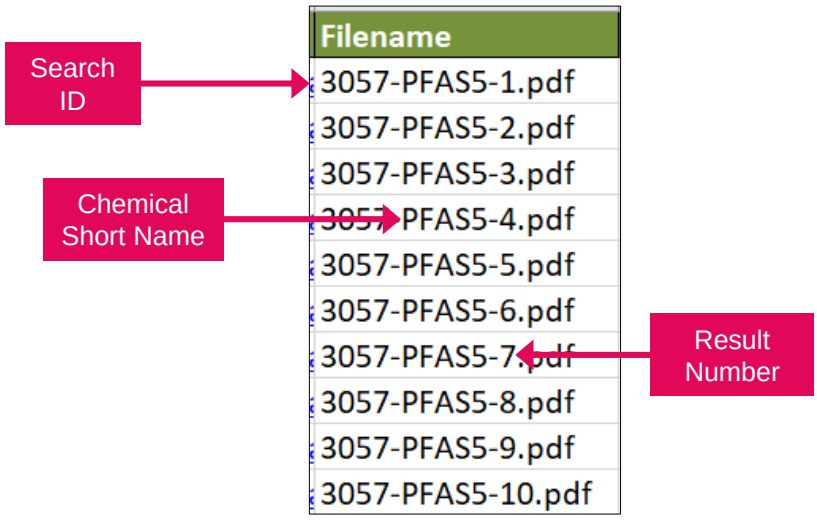
ECHA example

Capture citation information and relevant data

Step 3

Track!

Record filenames for each result



59

- Record each filename in the tracking sheet using an organized convention. Unique filename structures are not only useful for PDFs retrieved from ECHA but for all sources. These should match the actual filename that the document was saved under.
- The filename column will be used as a unique identifier for each result pulled into the tracking sheet. **This is especially important when HERO IDs are not yet assigned and you need to account for each unique result.**
- The following is an overview of what each section in the example filename shown here reflects (please keep in mind that this is a simplified version of filename):
    - **Search ID:** Unique code assigned to a specific source.
    - **Chemical Short Name:** Abbreviated chemical name.
    - **Result Number:** Sequential result number based on the order results were retrieved in. These refer to the specific entry/endpoint that data was extracted from and will only overlap in cases where multiple points of departure and data sources exist for each page.
- Depending on the source, more information can be added to the filename to identify where the reference was found. For example, in ECHA each dossier/nested-link pair may have its own identifier within a filename.

Capture citation information and relevant data

Step 3

Notes and best practices

- For sources considered to be **'live' data sources** (e.g., ECHA) that may be continuously updated, be sure to:
  - Extract the **'last modified' date**
  - Save each entry as a **PDF file** that can be referenced after extraction is complete, and enter the PDF filename in the tracking file
- ECHA-specific
  - Additional **references may be embedded** within sections of the dossier; these are not common occurrences but may happen more frequently under the 'Basic toxicokinetics' endpoint category.
  - Some dossiers will have **multiple entries** for the same endpoint depending on the number of submissions made; in these cases, each entry should be extracted as a separate row.
  - In some instances, multiple data sources and point-of-departure values are available within a dossier; each **unique combination** of data-source and point-of-departure **should be extracted as a separate row**. These rows will share the same general information and filename but differ in the values extracted for these sections.

Capture citation information and relevant data

Step 3

Track!

Tracking

- Extract:
  - Identification information for the results to be extracted
  - A link or reference back to the individual result
  - The PDF filename of the saved PDF for each record
  - Not all columns/data fields will be relevant for all rows, but all data fields of interest should be extracted into a separate column

Chemical-specific identifiers

Typically identical to Result.ID.Number and contains .pdf file extension

ChemShortname	DSSTox_Substance_iD	CAS Number	Chemical.Searched	Result.ID.Number	Search.Terms	Result.URL	Filename
1NAP	DTXSID6021793	90-15-3	1-naphthol		[CAS number]		

Abbreviated or short name for chemical

Unique naming convention for results

Link to the results pulled

61

Search process and tracking

Use this process to guide the search and track results:

Step 0: Consider project goals to determine plan-of-action and prepare tracking file (see previous section)

Step 1: Identify chemical search terms used for each source

Step 2: Perform initial search to determine what data exists in each source

Step 3: Capture citation information and relevant data

Step 4: De-duplicate and add new references to project database

De-duplicate and add new references to project database

Step 4

Managing and integrating gray literature

Accurate retrieval, tracking, and reference management are important to **identify duplicates** and create review-ready references.

1. Identify results **with** citation information
   - Search for existing HERO ID
   - Import the reference to HERO if the citation has no existing record
   - Deduplicate against other sources of gray literature and bibliographic database searches
2. Identify results **with little or no** citation information
   - Deduplicate based on available data fields from each source in the tracking file (e.g., for human health data, study details and results like PODs, endpoints, and species can be used in this process)
   - Manually create citations using preferred naming convention for each source
   - Import the citation to HERO to create a HERO record
3. Add any **confirmed new** references to the review steps (e.g., screening)

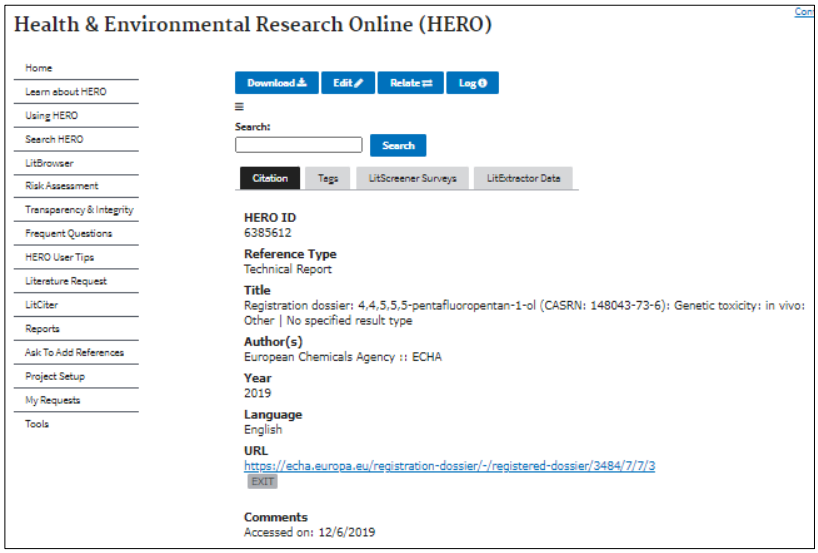
63

- Begin deduplicating results against citations from the peer-reviewed literature.
- Track the references that are already included in the review to prevent them from being added.
- Add only the confirmed ‘new’ references to the screening step.

De-duplicate and add new references to project database

Step 4

Gray-literature reference HERO entry



64

- This is an example of a complete HERO record for one of the PFAS results found in ECHA.
- Chemical-specific identifiers, endpoints, and result types are all included, as well as the date the information was last accessed on.
- It is important to keep in mind that, although URLs are provided for each entry, links often change due to updates made in the ECHA database.

De-duplicate and add new references to project database

Step 4

Notes and best practices

- The **'bibliographic source' information does not always reliably contain author, title, and year**; some sleuthing may be required to compile complete citations for gray-literature references
- It is useful to **record whether or not references have been tagged** to projects in HERO
- ECHA-specific:
  - For results that have multiple rows of data extracted for the same PDF with unknown citations, make sure to **account for these duplicates before creating HERO records**
  - If there is an entry with both **known and unknown** citations within the same ECHA dossier, the reference **with citation information** would be **assigned a HERO ID** and the reference **without citation information** would be **identified using the same HERO ID** as the cited reference. In many cases, the unnamed citations are secondary reference types and contain the same information that is in the cited reference.

De-duplicate and add new references to project database

Step 4

Notes and best practices

- It is important to remember that **URLs for gray-literature references may not be completely reliable** if the source contains 'live' data
  - For these types of sources, it is best to **rely on the actual PDFs** rather than the URLs to access the information
- ToxVal outputs can **help identify duplicates for references without citation information** by matching points-of-departure, species, sex, dose descriptor, endpoints, and other identifiable information.
  - It is possible to conduct comparisons between ToxVal and ECHA-only results by filtering the ToxVal output to only include ECHA IUCLID results

De-duplicate and add new references to project database

Step 4

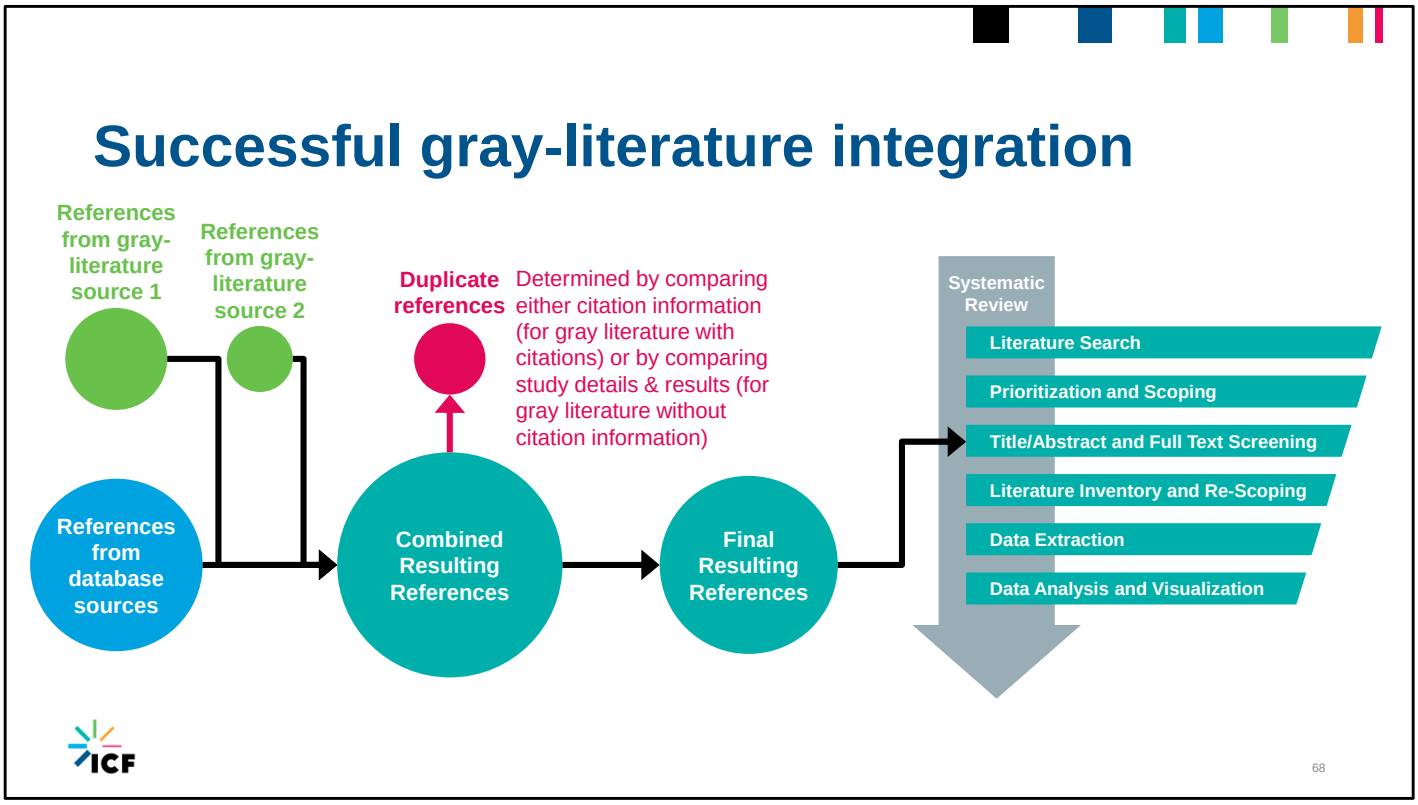
Track!

Tracking

- Concatenate any provided citation information into a column of the tracking sheet to begin to **form complete citations** that can be checked against HERO for possible duplicates or uploaded to HERO for new references.
- Record whether or not citation information was provided in a separate column of the tracking sheet to help **easily filter** between those that have citations and those that may need manual ones created.
- **Record HERO IDs of confirmed duplicate studies** in the gray-literature tracking file to maintain a record of duplicates that should not proceed into the pool of references for the project.

67

- Track the references that are already included in the review to prevent them from being added.
- Add only the confirmed ‘new’ references to the screening step.



- Completing the duplicate checks and incorporating the results of the gray-literature search concludes our training.
- We've added this visual to show how the gray-literature sources, in green, come into the project after the peer-review database sources.
    - The results are reviewed for duplicates against the peer-reviewed references and the final unique results move into the screening process.
- Questions can be sent to Parnian Soleymani and Nicole Vetter; the contract includes time for ICF to work with EPA on gray literature searches, so please reach out with questions.

- **Note:** The Appendix provides an example of a more detailed ECHA search.

Questions?

Questions

Direct questions about this presentation or the gray-literature search process to:

- **Nicole Vetter | ICF: nicole.vetter@icf.com**
- **Parnian Soleymani | ICF: parnian.soleymani@icf.com**

Detailed product examples

Examples of products incorporating gray literature

Other Sources Summary Table

- A table with complete information about the gray literature and other non-bibliographic database sources and the returned results.
- Depicts the total and unique results from each source.
- Included in the search strategy portion of a Systematic Review document.

Summary Table includes:

- Source searched
- Source location/ URL or other identifying information
- Terms used to search when applicable
- Date of search
- Total number of results
- Total number of unique results (included in the Study Flow Diagram)

Best Practices

- Keep a live tracking table that can be updated during the search, deduplication, and screening process.

6.4.11. Summary table for other sources search results

Source	Source Address	Search Terms	Search Date	Total Unique Number of Results Retrieved	Records not otherwise identified that were screened in DistillerSR
Review of reference lists studies considered relevant to PECO based on full-text screening.	NA	NA	NA	67	65
Review of reference lists from existing assessments (final or publicly available draft) or journal review articles that focused on human health	NA	NA	NA	3	0
EPA CompTox Chemicals Dashboard (ToxVal)	https://comptox.epa.gov/dashboard/dsstoxdb/results?abbreviation=TOXVAL_V5&search=DTXSID6021793#toxicity-values	90-15-3 (Results from Human Health: POD, Toxicity Value, Lethality Effect Level)	12/10/2019	21	6
ECHA	https://echa.europa.eu/information-on-chemicals/information-from-existing-substances-regulation	90-15-3	10/8/2019	53	24
EPA ChemView	https://chemview.epa.gov/chemview?tf=0&ch=90-15-3&su=2-5-6-7-37574985&as=3-10-9-8&ac=1-15-16-6378999&ma=4-11-1981377&tds=0&tdl=10&tas1=1&tas2=asc&tas3=undefined&tss=	90-15-3	9/19/2019	3	1
High Production Volume Challenge Database (HPVIS)	https://ofmpub.epa.gov/oppthpv/quicksearch.display?pChem=101850	90-15-3	9/19/2019	4	4

- The summary table includes an overview of the sources searched, the total number of unique references found in each source, and the number of references that were unique to the project.

Examples of products incorporating gray literature

Study Flow Diagram

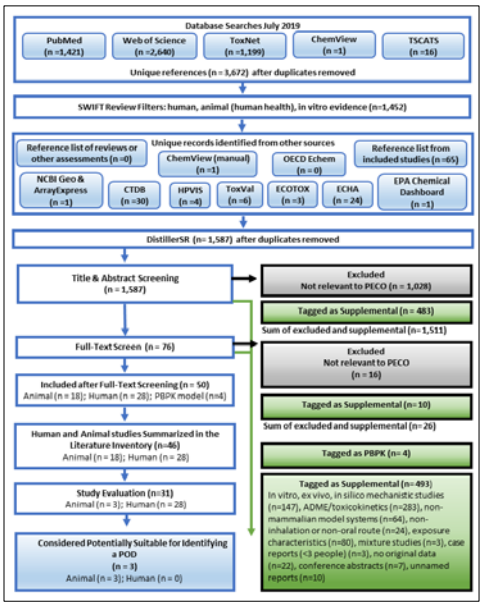
- **Also called Literature Flow Diagram or PRISMA Flow Diagram**
- **Provides a visual narrative of the flow of information through the different steps of the Systematic Review process**
- **Includes**
  - Records retrieved from the literature searches (peer & gray)
  - Machine learning or text analytic tools used
  - Include, exclude, and supplemental screening results

Gray Literature

- **Is generally searched concurrently or after the Peer literature**
- **Is brought into the Systematic Review process after any automation tools that rely on title/abstract are used**
- **Each gray-literature source and the number of results it brought to the review is identified**

Best Practices

- **Deduplicate the references found in gray sources against the peer literature results**
- **Include the unique result counts for each source**



73

- The study flow diagram expands on the number of references that were pulled into each step of the process and how they were categorized.

Example search process in ECHA for hazard-ID projects

Search process and tracking

Use this process to guide the search and track results:

Step 0: Consider project goals to determine plan-of-action and prepare tracking file

Step 1: Create a new tab for each source searched

Step 2: Identify the chemical search terms
(e.g., chemical name, CASRN, DTXSID, synonyms)

Step 3: Identify which search terms are used for each source

Step 4: Perform initial search to determine what data exists in each source

Step 5: Capture fields of interest to include in the tracking sheet

ICF

- Please note that the steps in this slide do not entirely match the process listed in the main slides as we have slightly expanded on these to provide additional details in the ECHA example.

Step 1:
Create a new tab for each source searched

Example ECHA Search

Track!

Create a new tab in the tracking sheet for ECHA

Incorporate the following **general** fields:

- Search ID
- Search Type
- Completed Date
- Has Results
- URL Searched

Search ID	Search Type	Completed Date	Has Results?	Url Searched
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances
3057	Manual	11/7/2019		https://echa.europa.eu/information-on-chemicals/registered-substances

ECHA_Results

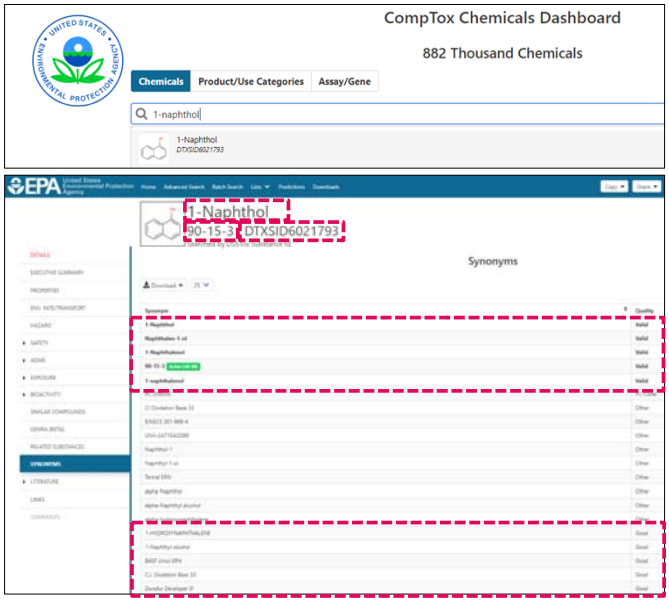
76

- The first step consists of creating a new tab for results that will be pulled from ECHA. The fields listed in this slide are typically incorporated in the tracking sheet for each source and contain information that is consistent across each chemical searched in ECHA.

Step 2:
Identify the chemical search terms

Example ECHA Search

Identify chemical terms using CompTox Dashboard



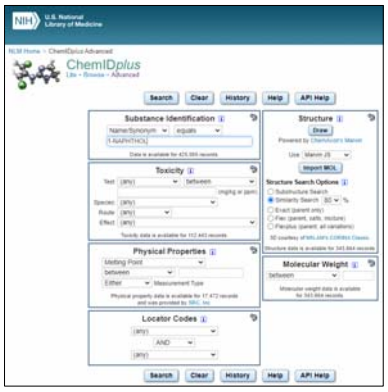
77

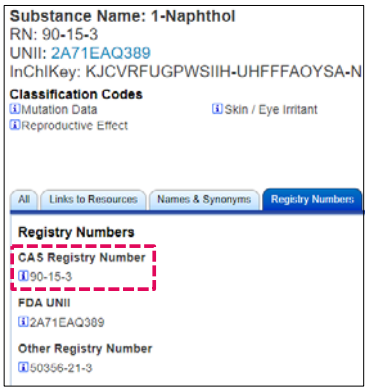
- The next step consists of identifying the chemical terms that will be used in the search process.
- The CompTox Dashboard can be searched to obtain the CAS number, DTXSID, and synonyms for each chemical.
- DTXSIDS are becoming more frequently utilized in chemical searches and are useful to incorporate in the tracking sheet.
- Depending on your search, you may want to limit the synonyms obtained through the CompTox Dashboard to those categorized as “valid”, “good”, and “fair”.
- In this example, the unique synonyms retrieved for 1-Naphthol from both ChemID*plus* and CompTox will be compiled.

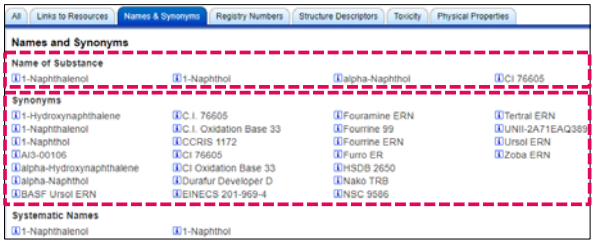
Step 2:
Identify the chemical search terms

Example ECHA Search

Identify chemical terms using ChemID*plus*







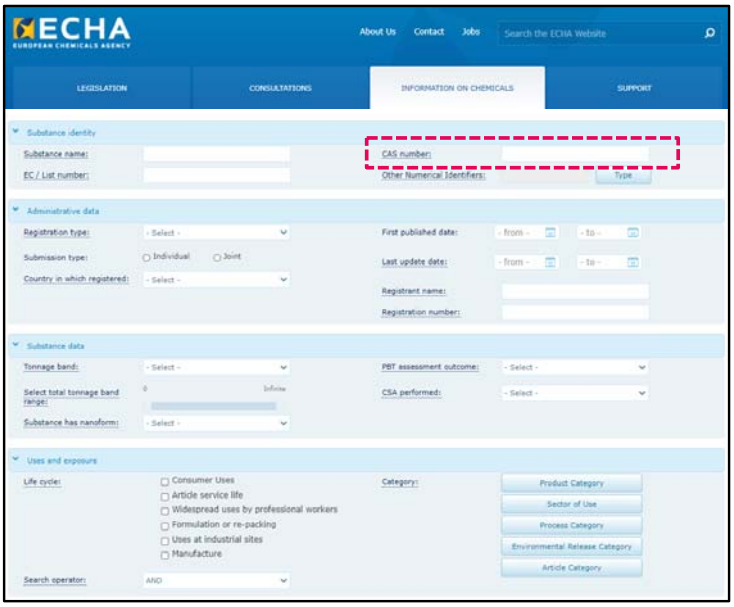
78

- The ChemID*plus* database can also be searched to obtain information on both the chemical CAS number and synonyms.
- The CAS number is typically used as the search term and the synonyms are more frequently utilized during the peer search and screening process. Therefore, you should keep a running list of these synonyms in a separate sheet for reference.
- The chemical identifiers found in the ChemID*plus* dashboard for 1-Naphthol are pictured above, with some of the more useful information outlined in red.

Step 3:
Identify which search terms are used for each source

Example ECHA Search

Identify the search terms used in ECHA database



79

- The third step involves identifying the specific search terms that will be used to search ECHA.
- As you can see, there are many different pieces of information you can use to filter and access chemical results in ECHA.
- ICF typically uses the CAS number field to retrieve results, but other identifiers can also be used to search the database.

Step 3:
Identify which search terms are used for each source

Example ECHA Search

Track!

Incorporate chemical columns in tracking sheet

Chemical-specific identifiers

Typically identical to Result.ID.Number and contains .pdf file extension

ChemShortname	DSSTox_Substance_iD	CAS Number	Chemical.Searched	Result.ID.Number	Search.Terms	Result.URL	Filename
1NAP	DTXSID6021793	90-15-3	1-naphthol		[CAS number]		

Abbreviated or short name for chemical

Unique naming convention for results

Link to the results pulled

80

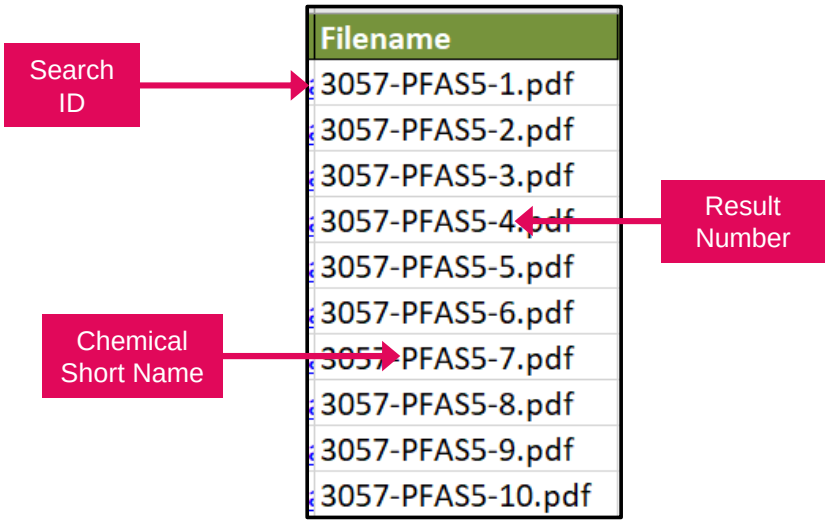
- At this point, you can incorporate the columns for chemical-specific identifiers and naming conventions used to track results.

Step 3:
Identify which search terms are used for each source

Example ECHA Search

Track!

Incorporate 'Filename' column in tracking sheet



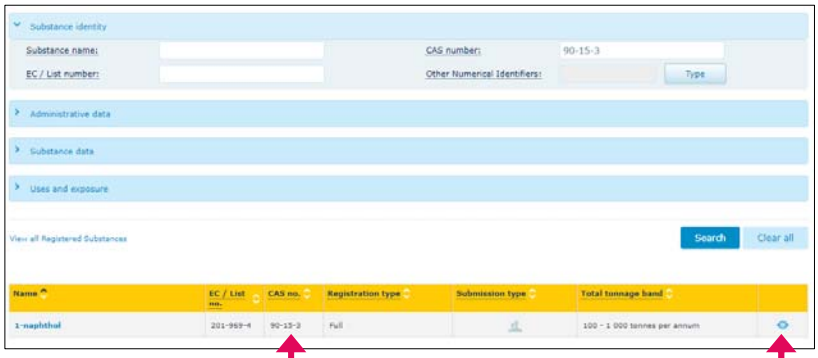
81

- The filename column will be used as a unique identifier for each result pulled into the tracking sheet. **This is especially important when HERO IDs are not yet assigned and you need to account for each unique result.**
- The following is an overview of what each section in the filename reflects:
    - **Search ID:** The unique code assigned to ECHA.
    - **Chemical Short Name:** The abbreviated chemical name for a given chemical.
    - **Result Number:** A sequential result number based on the order in which results were retrieved in. These refer to the specific entry/endpoint that data was extracted from and will only overlap in cases where multiple points of departure and data sources exist for each page. Additional numbers can be added to the filename to indicate the dossier/nested-link pair used to access the data.

Step 4:
Perform initial search to determine what data exists in each source

Example ECHA Search

Determine available data



Verify the CASRN under the CAS no. matches chemical of interest

Click on the "eye" icon to access the dossier

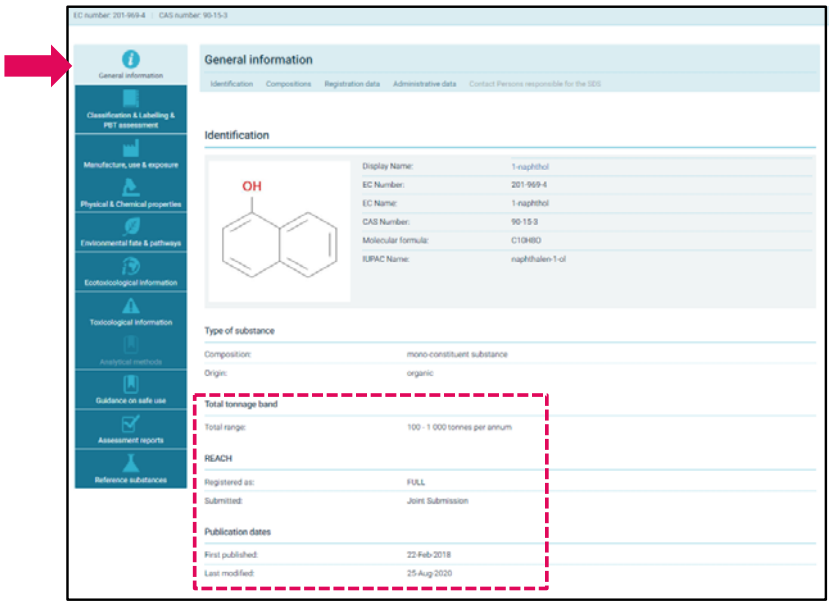
82

- The next step involves an initial search of the ECHA database to see what types of data exist.
- After entering the CASRN, results should populate below the search fields.
- Verify that the number under the CAS no. column matches the chemical you are searching. Sometimes different chemicals that are classified under the same chemical group can appear in the results. Therefore, it is important to confirm that the correct CAS number is listed before opening the dossier.
- In order to view the chemical dossier, click on the "eye" icon located to the far right.
- Some substances may have multiple results for different submission and registration types. It is important to capture information from all these dossiers.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Record relevant information



- Depending on the type of data you are interested in, you may choose to extract from specific tabs in the dossier.
- Information from the General Information tab is always extracted as it contains valuable data on the submission and type of data published.
- Since these dossiers often change and are considered to contain “live” data, it is important to capture the last modified date to indicate the exact date information was updated.
- Data extracted from this tab will be identical for all entries and rows within a specific chemical dossier submission.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Track!

Incorporate tab-specific columns in tracking sheet

1	2	3	4	5
Q	R	S	T	U
TotalTonnageBand_TotalRange	Reach_Registration	Reach_Submission	Publication_FirstSubmitted	Publication_LastUpdate

	Total tonnage band	
1	Total range:	100 - 1 000 tonnes per annum
	REACH	
2	Registered as:	FULL
3	Submitted:	Joint Submission
	Publication dates	
4	First published:	22-Feb-2018
5	Last modified:	25-Aug-2020

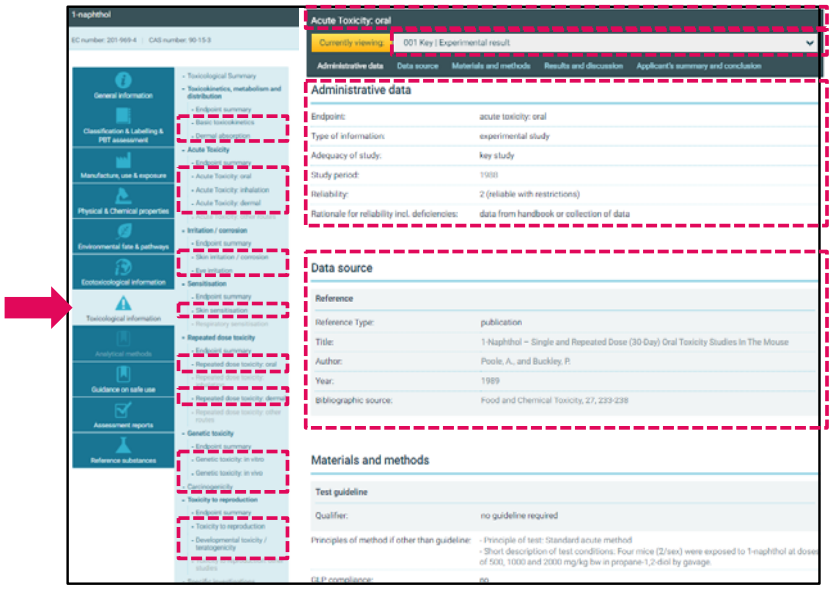
84

- Now that fields of interest are identified for the first tab, we can incorporate them in the tracking sheet.
- ICF recommends including the section name followed by the specific field when creating column headers since some field names may overlap between multiple sections. This also helps keep different data types grouped together.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Record relevant information

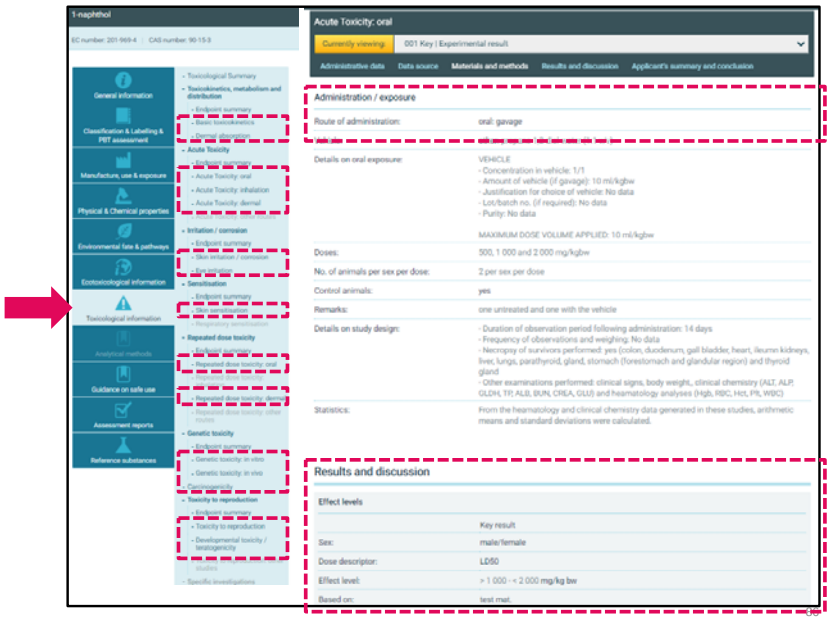


- Next, proceed to the next tab that has data you are interested in.
- For human health data, the Toxicological Information tab typically contains all the information we are interested in extracting.
- Each nested link under this tab would get its own row in the tracking sheet.
- We are not interested in summary pages since the information within these will be captured under the other endpoints mentioned for each category.
- Although the details you choose to extract may be unique based on the project, ICF typically pulls data in the fields outlined in this slide.
- Depending on the endpoint, there may be additional sections that are not captured here.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Record relevant information



- Since the types of information available may vary depending on the endpoint, it is important to explore different categories to see what types of information are provided under each.
- Some columns will only be utilized for specific endpoints whereas others will apply to all.

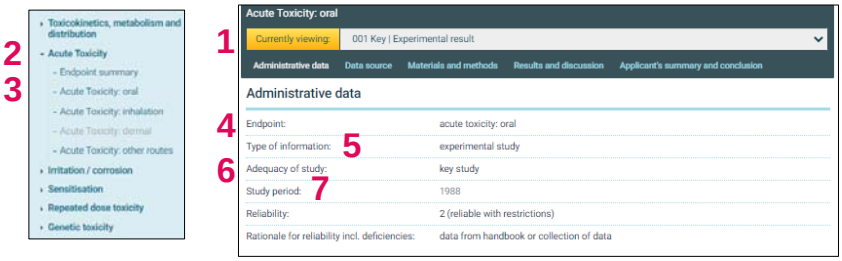
Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Track!

Incorporate tab-specific columns in tracking sheet

1	2	3	4	5	6	7
View.Name	ToxCategory	EndpointCategory	AdminData_Endpoint	AdminData_Type	AdminData_Adequacy	AdminData_StudyPeriod



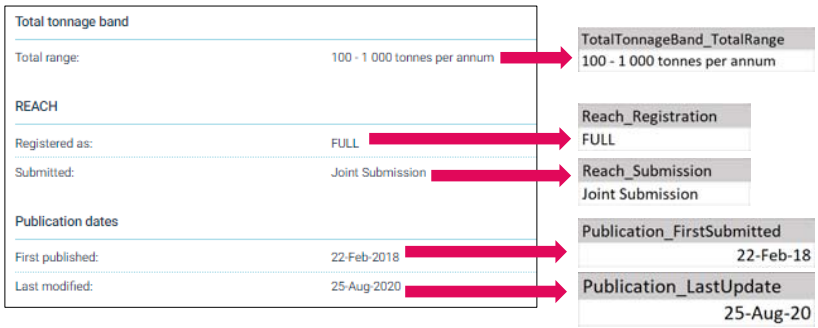
87

- As before, incorporate the tab-specific columns in the tracking sheet using an abbreviated section name followed by the specific field of interest.
- Note how the tox and endpoint categories are listed independently for each row.
    - The tox category is much broader and can be retrieved directly from the main menu which contains the nested endpoints.
    - Each individual link acts as a subsection for specific endpoint categories.
    - By breaking these elements up, it is easier to filter for general and specific endpoints as needed.
- The View.Name column is also included to reflect the name provided in the drop-down menu for each entry. This information will make it easier to create HERO citations and refer to specific entries if data is updated in the ECHA database.
- The columns shown in this slide are only a few out of the many that will be included in the tracking sheet.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Extract relevant data



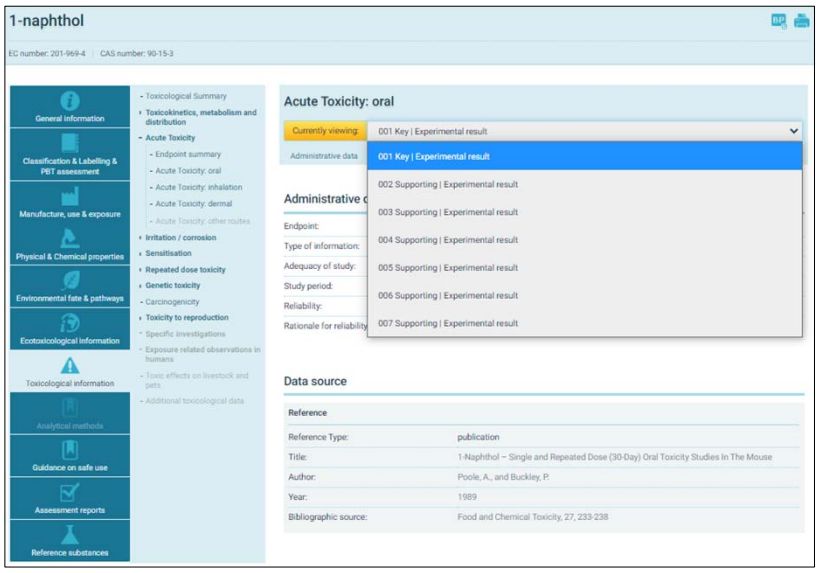
88

- Once all fields of interest are identified and incorporated in the tracking sheet, you may begin extracting information from each section in the dossier.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Extract multiple entries into multiple rows

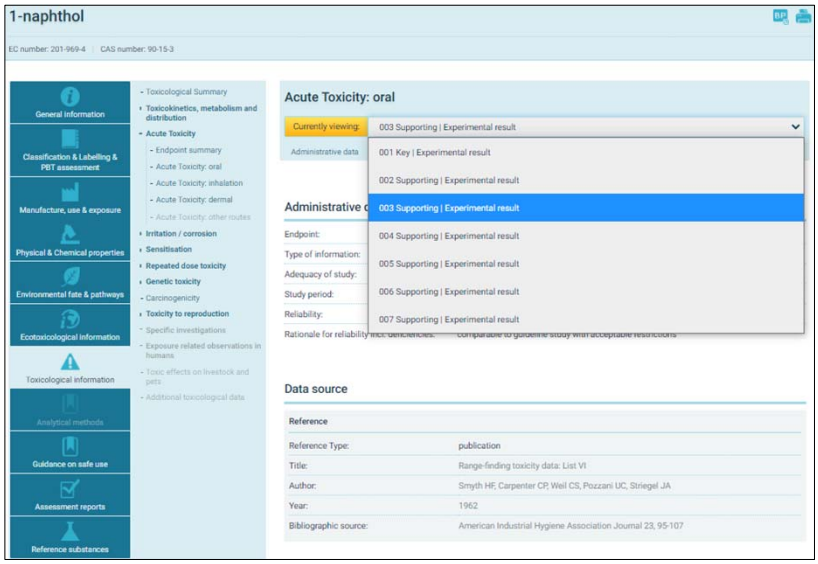


- Some dossiers will have multiple entries for the same endpoint depending on the number of submissions made.
- We want to account for all entries as they typically contain different data. Therefore, each entry should be extracted as a separate row with the name provided in the drop-down menu extracted under the View.Name column.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Extract multiple entries into multiple rows

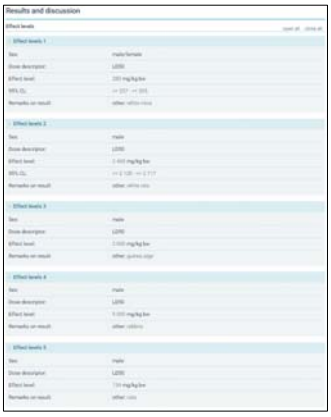


- Notice how the information provided on each page changes as you select a different entry in the drop-down menu.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Extract multiple data sources/PODs into multiple rows



DataSource_ReferenceType	DataSource_Title	DataSource_Author	DataSource_Year	DataSource_ReportDate	DataSource_BibliographicSource	Results_Sex	Results_DoseDescriptor	Results_EffectLevel
publication	EXPERIMENTAL MATERIALS F	Rakhmanin, Yu. A	1969		Rakhmanin, Yu. A.; Promyshlennye	male/female	LD50	280 mg/kg bw
publication	EXPERIMENTAL MATERIALS F	Rakhmanin, Yu. A	1969		Rakhmanin, Yu. A.; Promyshlennye	male	LD50	2 400 mg/kg bw
publication	EXPERIMENTAL MATERIALS F	Rakhmanin, Yu. A	1969		Rakhmanin, Yu. A.; Promyshlennye	male	LD50	2 000 mg/kg bw
publication	EXPERIMENTAL MATERIALS F	Rakhmanin, Yu. A	1969		Rakhmanin, Yu. A.; Promyshlennye	male	LD50	9 000 mg/kg bw
publication	EXPERIMENTAL MATERIALS F	Rakhmanin, Yu. A	1969		Rakhmanin, Yu. A.; Promyshlennye	male	LD50	134 mg/kg bw
review article or handbook	Handbook of Environmental D	Verschueren, K	2001		Verschueren, K. Handbook of Envir	male/female	LD50	280 mg/kg bw
review article or handbook	Handbook of Environmental D	Verschueren, K	2001		Verschueren, K. Handbook of Envir	male	LD50	2 400 mg/kg bw
review article or handbook	Handbook of Environmental D	Verschueren, K	2001		Verschueren, K. Handbook of Envir	male	LD50	2 000 mg/kg bw
review article or handbook	Handbook of Environmental D	Verschueren, K	2001		Verschueren, K. Handbook of Envir	male	LD50	9 000 mg/kg bw
review article or handbook	Handbook of Environmental D	Verschueren, K	2001		Verschueren, K. Handbook of Envir	male	LD50	134 mg/kg bw
review article or handbook	Sax's Dangerous Properties of	Lewis, R.J. Sr	2004		Lewis, R.J. Sr (ed) Sax's Dangerou	male/female	LD50	280 mg/kg bw
review article or handbook	Sax's Dangerous Properties of	Lewis, R.J. Sr	2004		Lewis, R.J. Sr (ed) Sax's Dangerou	male	LD50	2 400 mg/kg bw
review article or handbook	Sax's Dangerous Properties of	Lewis, R.J. Sr	2004		Lewis, R.J. Sr (ed) Sax's Dangerou	male	LD50	2 000 mg/kg bw
review article or handbook	Sax's Dangerous Properties of	Lewis, R.J. Sr	2004		Lewis, R.J. Sr (ed) Sax's Dangerou	male	LD50	9 000 mg/kg bw
review article or handbook	Sax's Dangerous Properties of	Lewis, R.J. Sr	2004		Lewis, R.J. Sr (ed) Sax's Dangerou	male	LD50	134 mg/kg bw

91

- In some instances, there will be multiple data sources and point of departure values available within a dossier.
- Each data source and effect level section should be extracted as a separate row.
- These rows will share the same filename and information across other sections but differ in the values extracted for the data source and effect levels.
- This slide depicts how the combination of three different references listed under the data source section and five effect levels were extracted as 15 rows.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Look for additional references in different sections

Results and discussion

Toxicokinetic / pharmacokinetic studies

Details on absorption:	Please see the excutive summary on oral, dermal and inhalatory absorption of 1-naphthol.
Details on distribution in tissues:	Please see the excutive summary
Details on excretion:	Please see the excutive summary

Metabolite characterisation studies

Details on metabolites:	Please see the excutive summary

Bioaccessibility

Bioaccessibility testing results:	Conclusion: Bioavailable

Any other information on results incl. tables

References cited:

Best, E. M. Jr., Murray, B. L.,J. Occupational .Med.4,507, 1962.

Carpenter, C. P., Weil, C. S., Palm. P. E.. Woodside, M. W.,Nair, J. H. III, Smyth, H. J., Jr.: J. Agr. Food Chem. 9, 30, 1961.

Doherty at al, Biochem Pharmacol 33(4): 543-9, 1984.

Doherty M D,Cohen GM,Gant TW,Naish S,Riley PA.Metabolism of 1-naphthol by tyrosinase.BiochemPharmacol.34(17):3167-72,1985.

Dechelotte, P., Varrentrapp, M., Meyer, H.J., and Schwenk, M. Conjugation of 1-naphthol in human gastric epithelial cells. Gut 34 (2): 177-80, 1993.

Knaak. J. B.. Tallant, M.J., Bartley, W. J., Sullivan L. J.: J. Agr. Food Chem. 13, 537, 1965.

Leeling, N.C., and Casida, JE.: Metabolites of Carbaryl (1-Naphthyl Methylcarbamate) in Mammals and Enzymatic Systems for Their Formation. J. Agr. Food Chem. Vol. 14 (3) 281-290, 1966.

Mehta R,Hirom PC,Millburn P. The influence of dose on the pattern of conjugation of phenol and 1-naphthol in non-human primates. Xenobiotica, 8(7): 445-52, 1978

P. Suwan-ampai, A. Navas-Acien, P.T. Strickland, and J. Agnew: Involuntary Tobacco Smoke Exposure and Urinary Levels of Polycyclic Aromatic Hydrocarbons in the United States, 1999 to 2002. Cancer Epidemiol Biomarkers Prev.:18(3)884-93, 2009.

Whitehurst, W.E., Bishop, E.T., Critchfield, F. E., Gyrisco, G. G., Huddleston. E. E.W.,Arnold, H., Lisk, D. L.: J. Agr. Food Chem., 11, 167, 1963.

Note: Other citations can be found at various endpoint summaries of this dossier.

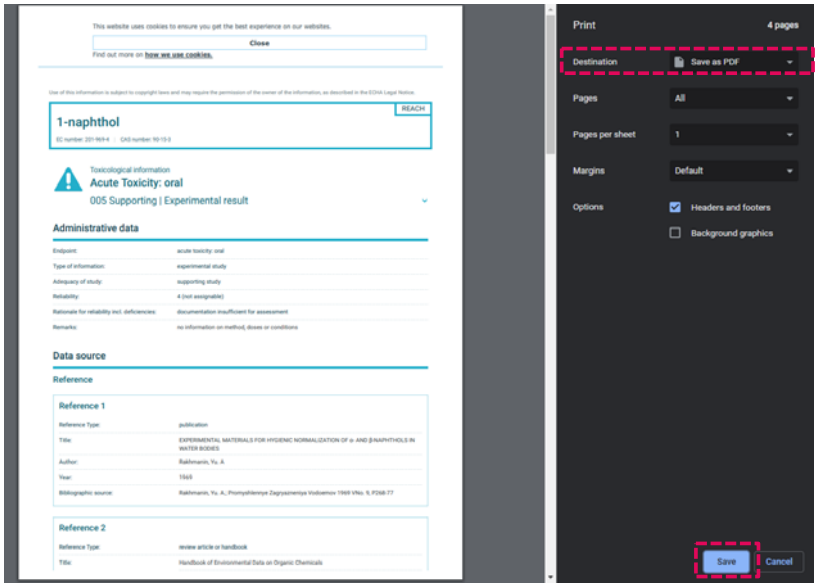
92

- Look out for and extract additional references that may be embedded within different sections of the dossier.
- These are not common occurrences but may happen more frequently under the Basic Toxicokinetics endpoint category.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Save dossiers as PDFs



93

- Once extraction is complete, save each entry as an individual PDF.
- You can click on the “print” icon at the top of the page and save each file using the naming convention under the filename column.

Step 5:
Capture fields of interest to include in the tracking sheet

Example ECHA Search

Track!

Tracking

- ECHA-specific:
  - Some substances may have multiple results for different submission and registration types; it is important to **capture information from all dossiers** of the chemical of interest.
  - ICF recommends column-header naming convention of **including the tab or area of the page + specific data** being extracted to help keep columns organized (e.g., 'Data Source_Reference Type', 'Publication_Last Update')
  - The **'View.Name' column** is recorded to reflect the name provided in the 'Currently Viewing' drop-down menu for each entry. This information will make it **easier to create HERO citations** and refer to specific entries if data are updated in the ECHA database.
  - In some cases, **multiple endpoints** may belong to the **same unnamed study** as indicated by the points of departure, guidelines, and other information extracted. To detect where overlaps exist, **concatenate data** found across chemical, study period, citations information, reliability, and point of departure columns, and apply conditional formatting to detect duplicate entries.
    - References that share the same information, including POD values and specific guidelines mentioned under the reliability column, will be grouped and assigned one HERO ID.

94

- It is important to keep these points in mind as you go through the process.

Using the results

Example ECHA Search

How to utilize results

Reference Management: Once searches have been completed, results downloaded or extracted, and unique identifiers have been assigned, reference management begins. Accurate retrieval, tracking, and reference management are important to identify duplicates and create references ready for screening.

Step 1: Identify results with citation information

- Search for existing HERO ID and import the reference to HERO if the citation has no existing record
- Deduplicate against other sources of gray literature and bibliographic database searches

Step 2: Identify results with little or no citation information

- Deduplicate based on available data fields from each source in the search tracking file
- Manually create citations using preferred naming convention for each source and import the citation to HERO to create a HERO record

Step 3: Add the results to screening tool

95

- Once the search process is complete and all the data has been compiled in the tracking sheet, reference management begins.
- There are three main steps used to identify, categorize, and account for results.
    - Those with citation information will be searched in HERO and deduplicated against other sources and searches. If records in HERO do not exist, one will be created for the reference.
    - Results with little or no citation information will be deduplicated based on data retrieved from other sources. Those that are new will have unique citations created based on the ECHA entry and will be imported to HERO as a new record.
    - All results that are unique to the project and have not been yet been accounted for will be added to the screening process.

Using the results

Example ECHA Search

Assign HERO IDs

DataSource_Title	DataSource_Author	DataSource_Year	DataSource_BibliographicSource	Citation?	Citation Information	Multiple Citations?	HEROID
Beitrag zur Toxizi	Schmidt P	1980	Zeitschrift für die gesamte Hygier	Yes	Schmidt P. Schmidt P. Beit		
Beitrag zur Toxizi	Schmidt P	1980	Zeitschrift für die gesamte Hygier	Yes	Schmidt P. Schmidt P. Beit	same	
Safety Data Sheet	Apollo Scientific Limited		Safety Data Sheet of Appollo Scie	Yes	Apollo Scientific Limited. .		
Beitrag zur Toxizi	Schmidt P	1980	Zeitschrift für die gesamte Hygier	Yes	idt P. Schmidt P. Beit		
Beitrag zur Toxizi	Schmidt P	1980	Zeitschrift für die gesamte Hygier	Yes	hmidt P. Schmidt P. Beit		
Safety Data Sheet	Apollo Scientific Limited		Safety Data Sheet of Appollo Scie	Yes	Apollo Scientific Limited. .		
Unnamed		2019		Not Provided	Unnamed, 2019		
Beitrag zur Toxizi	Schmidt P	1980	Zeitschrift für die gesamte Hygier	Yes	Schmidt P. Schmidt P. Beit		
Safety Data Sheet	Apollo Scientific Limited		Safety Data Sheet of Appollo Scie	Yes	Apollo Scientific Limited. .		
Unnamed		2019		Not Provided	Unnamed, 2019		
Unnamed		2019		Not Provided	Unnamed, 2019		
Unnamed		2002		Not Provided	Unnamed, 2002		
Unnamed		2007		Not Provided	Unnamed, 2007		
Unnamed		2015		Not Provided	Unnamed, 2015		
Unnamed		2007		Not Provided	Unnamed, 2007		

Citation Information
Apollo Scientific Limited. Apollo Scientific Limited. Safety Data Sheet 1H,1H,7H-Dodecafluoroheptan-1-ol 98%. Safety Data Sheet of Appollo Scientific Limited, Manchester, UK; Revision date 2012-09-14

HERO ID
6301178

Reference Type
Technical Report

Subtype
Fact Sheet

Title
Safety Data Sheet: 1H,1H,7H-Dodecafluoroheptan-1-ol 98%

Author(s)
Apollo Scientific Limited :: Apollo Scientific

Year
2012

Publisher
Apollo Scientific Limited

Location
Manchester, UK

96

- Additional columns will be incorporated which will help identify, develop, and assign HERO IDs to each unique ECHA result.
- For results with citation information provided, concatenate the Title, Author, and Year columns into a Citation Information column. In some cases, information from the Bibliographic Source field will also need to be included if a specific journal, page, and other identifiable pieces are mentioned. The reason that we do not solely pull information from the Bibliographic Source column is because these are often incomplete and do not always include the author, title, and year.
- For results without citation information (e.g., unnamed studies), concatenate the Title and Year columns under the Citation Information column as before. The title is typically always “unnamed” for these studies; however, in cases where that is not mentioned or blank you can manually add that.
- Mark whether actual citations were provided under the data source section under the Citation? column. These will filter as “Yes” for rows with citation information and “Not Provided” for those with limited or unnamed studies.
- In cases where multiple rows exist for the same entry due to more than one point of departure or data source mentioned, these rows are indicated with “same” under the Multiple Citations? column so that they are not accounted for twice.
- Once all the citation information is compiled, you may begin assigning or creating HERO IDs for those marked as “Yes” under the Citation? column. Those marked as “Not Provided” will need to be treated differently as seen on the following slides.

Using the results

Example ECHA Search

Assign HERO IDs

HEROID	Tagged to Project	Found in Peer Search
6301181	No	No
6301181	No	No
6301178	No	No
6301181	No	No
6301181	No	No
6301178	No	No
6302585	No	No
6301181	No	No
6301178	No	No
6302586	No	No
6305768	No	No
6305770	No	No

97

- After HERO IDs have been assigned to results with citation information, check to see if any of these have been captured in the peer search. Those that are new will need to be added.
- It is also useful to track whether references have been tagged to projects in HERO.

Using the results

Example ECHA Search

Deduplicate unnamed studies

toxval_type_original	toxval_numeric	toxval_units	sex	critical_effect	details_text
NOAEL	25	mg/kg-	male	fertility	ECHA IUCLID

DataSource	Results	Results	Results	Results
Unnamed	male	NOAEL	25 mg/kg	fertility

98

- Since unnamed studies have limited citation information, it is important to conduct comparisons against other sources such as ToxVal to determine whether any matches can be made based on the points of departure, species, sex, dose descriptor, endpoints, and other identifiable information.
- By filtering the ToxVal output to only include ECHA IUCLID results, you can conduct comparisons against results retrieved from ECHA.
- Unnamed studies that cannot be matched will have HERO IDs created, whereas those that are duplicates of ToxVal results will be assigned the same HERO ID.

Using the results

Example ECHA Search

Assign HERO IDs

ToxCategory	EndpointCategory	AdminData_Type	AdminData_Adequacy	HERO Title	HERO ID
Genetic toxici	Genetic toxicity: in vitro		other information	Registration dossier: 4,4,5,5,5-pentafluoropentan-1-ol (CAS/RN: 148043-73-4) | Genetic toxicity: in vitro | 002 Other | No specified result type	6385610
Genetic toxici	Genetic toxicity: in vivo		other information	Registration dossier: [illegible]	6385612
Sensitisation	Skin sensitisation	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Skin sensitisation: 001 Key | Experiment	6385613
Repeated dos	Repeated dose tox	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Repeated dose toxicity: oral: Key | Expe	6385614
Repeated dos	Repeated dose tox	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Repeated dose toxicity: inhalation: Key	6385615
Genetic toxici	Genetic toxicity: in	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Genetic toxicity: in vitro: 001 Key | Expe	6385617
Genetic toxici	Genetic toxicity: in	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Genetic toxicity: in vitro: 002 Key | Expe	6385618
Genetic toxici	Genetic toxicity: in	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Genetic toxicity: in vitro: 003 Key | Expe	6385619
Toxicity to rep	Toxicity to reprod	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Toxicity to reproduction: Key | Experime	6385620
Acute Toxicity	Acute Toxicity: ora	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Acute Toxicity: oral: Key | Experimental	6385621
Acute Toxicity	Acute Toxicity: der	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Acute Toxicity: dermal: Key | Experiment	6385625
Irritation / cor	Skin irritation/corr	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Skin irritation/corrosion: 001 Key | Expe	6385626
Irritation / cor	Skin irritation/corr	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Skin irritation/corrosion: 002 Key | Expe	6385628
Irritation / cor	Eye irritation	experimental stuc	key study	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Eye irritation: 001 Key | Experimental re	6385629
Toxicokinetic	Basic toxicokinetics		other information	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Basic toxicokinetics: Other | No specifie	6385630
Acute Toxicity	Acute Toxicity: oral		other information	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Acute Toxicity: oral: Other | No specifie	6385631
Acute Toxicity	Acute Toxicity: dermal		other information	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Acute Toxicity: dermal: Other | No spec	6385633
Irritation / cor	Skin irritation/corrosion		other information	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Skin irritation/corrosion: Other | No spe	6385635
Irritation / cor	Eye irritation		other information	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Eye irritation: Other | No specified resul	6385636
Sensitisation	Skin sensitisation		other information	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Skin sensitisation: Other | No specified	6385637
Repeated dos	Repeated dose toxicity: oral		other information	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Repeated dose toxicity: oral: Other | No	6385638
Genetic toxici	Genetic toxicity: in vitro		other information	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Genetic toxicity: in vitro: 001 Other | No	6385639
Genetic toxici	Genetic toxicity: in vitro		other information	Registration dossier: 3-ethoxy-1,1,1,2,3,4,4,5,5,6,6,6-dodecafluoro-2-(trifluoromethyl)-hexane (CAS/RN: 297730-93-9): Genetic toxicity: in vitro: 002 Other | No	6385641
Irritation / cor	Skin irritation/corr	experimental stuc	weight of evidence	Registration dossier: 2,2,3,3,4,4,5,5-octafluoropentyl methacrylate (CAS/RN: 355-93-1): Skin irritation/corrosion: 001 Weight of evidence | Experimental result	6385642
Irritation / cor	Skin irritation/corr	experimental stuc	weight of evidence	Registration dossier: 2,2,3,3,4,4,5,5-octafluoropentyl methacrylate (CAS/RN: 355-93-1): Skin irritation/corrosion: 002 Weight of evidence | Experimental result	6385643
Irritation / cor	Skin irritation/corr	experimental stuc	weight of evidence	Registration dossier: 2,2,3,3,4,4,5,5-octafluoropentyl methacrylate (CAS/RN: 355-93-1): Skin irritation/corrosion: 003 Weight of evidence | Experimental result	6385644
Irritation / cor	Skin irritation/corr	experimental stuc	weight of evidence	Registration dossier: 2,2,3,3,4,4,5,5-octafluoropentyl methacrylate (CAS/RN: 355-93-1): Skin irritation/corrosion: 004 Weight of evidence | Experimental result	6385655
Irritation / cor	Eye irritation	experimental stuc	key study	Registration dossier: 2,2,3,3,4,4,5,5-octafluoropentyl methacrylate (CAS/RN: 355-93-1): Eye irritation: 001 Key | Experimental result	6385656
Sensitisation	Skin sensitisation	experimental stuc	key study	Registration dossier: 2,2,3,3,4,4,5,5-octafluoropentyl methacrylate (CAS/RN: 355-93-1): Skin sensitisation: 001 Key | Experimental result	6385658
Repeated dos	Repeated dose tox	experimental stuc	supporting study	Registration dossier: 2,2,3,3,4,4,5,5-octafluoropentyl methacrylate (CAS/RN: 355-93-1): Repeated dose toxicity: oral: Supporting | Experimental result	6385659

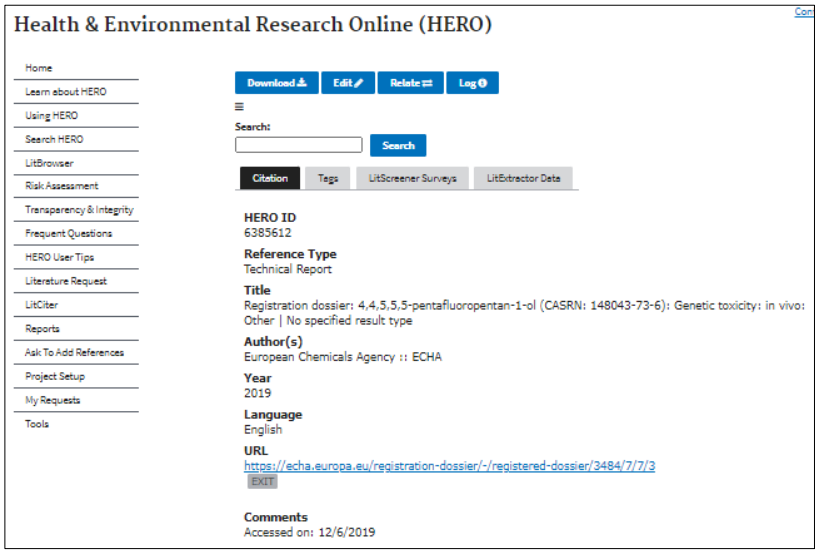
99

- For unnamed studies, HERO records will contain the same pieces of information and follow similar conventions.
- You will need to create unique citations based on the information extracted for each result.
- These pieces will be combined into the title column and incorporated in the HERO record.

Using the results

Example ECHA Search

Assign HERO IDs



100

- This is an example of a complete HERO record for one of the PFAS results found in ECHA.
- Chemical-specific identifiers, endpoints, and result types are all included, as well as the date the information was last accessed on.
- It is important to keep in mind that, although URLs are provided for each entry, links often change due to updates made in the ECHA database.

Example ECHA Search

Source-specific challenges

- Dealing with “live” data
- Dossiers with unnamed studies
  - No reference information
  - Detecting duplicates using points of departure, endpoints, species, and other identifiable information
  - Manual creation of unique citations
- Downloading files
- Multiple entries for a given endpoint

101

- With “live” data, it is especially important to incorporate the access date in both the tracking sheet and HERO record.
    - As mentioned in the previous slide, changes are constantly being made to the ECHA database. Therefore, tracking this information may help identify what types of data could be outdated.
    - The access date can be compared to the “last modified” date under the General Information tab of the dossier to determine whether updates have been made since extraction.
    - Despite this being a useful measure, it can still be somewhat challenging to stay on top of the most up-to-date data.
- To download PDFs, you must manually save and name each entry using the filename convention.

Example Searches in Other Sources

Prepare for searching other sources following the same steps:

Step 1: **Create a new tab for each source searched**

Step 2: **Identify the chemical search terms** (e.g., chemical name, CASRN, DTXSID, synonyms)

Step 3: **Identify which search terms are used for each source**

Step 4: **Perform initial search to determine what data exists in each source**

Step 5: **Capture fields of interest to include in the tracking sheet**

102

- This slide serves as a reminder of the general steps used in the search process.
- After completing the ECHA search, follow the same steps for each of the other sources.
- Each source will have slightly different search requirements, fields of interest, and data output.
- Since chemical search terms have already been identified prior to the ECHA search, you can skip this step for the remaining sources.

Step 3:
Identify which search terms are used for each source

Example Searches in Other Sources

ECHA

- Substance Identity
  - Substance Name
  - EC/List Number
  - CAS Number
  - Other Numerical Identifiers
- Administrative Data
- Substance Data
- Uses and Exposure

ChemView

- Chemical Name or Chemical Identifier
- Use
- Significant New Use Notification
- EPA Chemical Lists
- Chemical Category
- Effects/Endpoints
- Output Selections
  - Information Submitted to EPA, EPA Assessments, EPA Actions, Manufacturing, Processing, Use, and Release Data Maintained by EPA

NTP Database

- CAS Number
- Chemical Name
- Publication
- Project
- Study Number
- Study and Data Types
- NTP Data Only (filter)

OECD SIDS

- Chemical Identifier
- Types
  - Property Information
  - Exposure and Use Information
  - GHS Classifications
- Data Sources
  - OECD HPV, OECD SIDS IUCLID, SIDS UNEP*

103

- As before, identify the search terms for each source/database. Some sources may be limited to searching using the CASRN, DTXSID, or chemical name.
- Syntax may matter depending on the database searched (e.g., CEBS).
- ICF recommends applying the search terms and categories indicated in red since this is what is typically used to search each source.
- *For OECD SIDS, there are many more data sources available to choose from but those currently listed are ones ICF had previously identified as being of interest.

Using the results

Example Searches in Other Sources

How to utilize results

Step 1: Identify results with citation information

- Search for existing HERO ID and import the reference to HERO if the citation has no existing record
- Deduplicate against other sources of gray literature and bibliographic database searches

Step 2: Identify results with little or no citation information

- Deduplicate based on available data fields from each source in the search tracking file
- Manually create citations using preferred naming convention for each source and import the citation to HERO to create a HERO record

Step 3: Add the results to screening tool

104

- The same process of reference management applies once results have been collected from each source.

Presenting the results

Example Searches in Other Sources

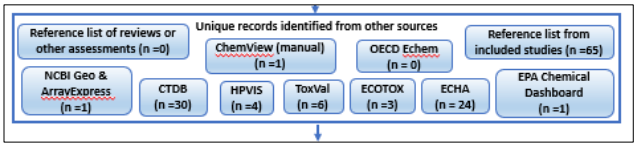
Completed Study Flow and Summary Table:

- Summary Table
  - Include search and result details for each source

6.4.11. Summary table for other sources search results

Source	Source Address	Search Terms	Search Date	Total Unique Number of Results Retrieved	Records not otherwise identified that were screened in (Distiller SR)
Review of reference lists studies considered relevant to PECO based on full-text screening	NA	NA	NA	67	65
Review of reference lists from existing assessments (final or publicly available draft) or journal review articles that focused on human health	NA	NA	NA	3	0
EPA CompTox Chemicals Dashboard (ToxVal)	https://comptox.epa.gov/dashboard/dashsearch/results?abbreviation=TOXVAL_V5&search=DTXSID6021793#toxicity-values	90-15-3 (Results from Human Health: POD, Toxicity Value, Lethality Effect Level)	12/10/2019	21	6
ECHA	https://echa.europa.eu/information-on-chemicals/information-from-existing-substances-regulation	90-15-3	10/8/2019	53	24
EPA ChemView	https://chemview.epa.gov/chemview?sf=0&ch=90-15-3&au=2-5-6-7-37574985&ae=3-10-9-8&ac=1-15-16-61789958&mr=8-11-19811777&nte=0&tdi=10&tas1=1&tas2=and&tas3=and&fixed&tsy=	90-15-3	9/19/2019	3	1
High Production Volume Challenge Database (HPVIS)	https://ofmpub.epa.gov/oppthpr/quicksearch.display?pChem=101850	90-15-3	9/19/2019	4	4

- Study Flow
  - Include each source with unique results



105

- Unique results will be incorporated in both the summary table and study flow diagram to show search and result details for all sources.