EPA/635/R-19/049
IRIS Assessments Protocol
www.epa.gov/iris

Systematic Review Protocol for the PFBA, PFHxA, PFHxS, PFNA, and PFDA IRIS Assessments

CASRN 335-76-2 (PFDA)
CASRN 375-95-1 (PFNA)
CASRN 307-24-4 (PFHxA)
CASRN 355-46-4 (PFHxS)
CASRN 375-22-4 (PFBA)

October 2019

Integrated Risk Information System
Center for Public Health and Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC

DISCLAIMER

This document is a public comment draft. This information is distributed solely for review purposes under applicable information quality guidelines. It has not been formally disseminated by EPA. It does not represent and should not be construed to represent any Agency determination or policy. It is being circulated for comments on its technical clarity and science policy implications. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

CONTENTS

TABLES

FIGURES

ABBREVIATIONS

ADME	absorption, distribution, metabolism, and excretion
AFFF	aqueous film-forming foam
AK DEC	Alaska Department of Environmental Conservation
ALT	alanine aminotransferase
AOP	adverse outcome pathway
AST	aspartate transaminase
ATSDR	Agency for Toxic Substances and Disease Registry
BMDL	benchmark dose lower confidence limit
BMI	body mass index
BMR	benchmark response
$BW^{3/4}$	body-weight scaling to the 3/4 power
CAR	constitutive androstane receptor
CAS	Chemical Abstracts Service
CASRN	Chemical Abstracts Service registry number
CBI	confidential business information
CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act
CLA	clearance in animals
CLH	clearance in humans
CPAD	Chemical and Pollutant Assessment Division
CPHEA	Center for Public Health and Environmental Assessment
CPN	chronic progressive nephropathy
CRD	chemical reporting data
CT DPH	Connecticut Department of Health
CTL	cytotoxic T lymphocyte
CWA	Clean Water Act
DNA	deoxyribonucleic acid
DTH	delayed-type hypersensitivity
DWEL	drinking water equivalent level
ECHA	European Chemicals Agency
EFSA	European Food Safety Authority
EPA	Environmental Protection Agency
FDA	Food and Drug Agency
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
FOB	functional operational battery
FXR	farnesoid X receptor
GLP	good laboratory practices
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HA	health advisory
HAWC	Health Assessment Workspace Collaborative
HED	human equivalent dose
HERO	Health and Environmental Research Online
HFPO	hexafluoropropylene oxide
hPPARα	humanized peroxisome proliferator-activated receptor alpha
HRL	health risk limit
i.p.	intraperitoneal
IARC	International Agency for Research on Cancer
IPCS	International Programme on Chemical Safety
IRIS	Integrated Risk Information System
IUR	inhalation unit risk
K	potassium
LD_{50}	median lethal dose
LOAEL	lowest-observed-adverse-effect level
LOD	limit of detection
MAC	maximum acceptable concentration
MCL	maximum contaminant level
MDH	Minnesota Department of Health
MF	modifying factor
MLR	mixed leukocyte reaction
MOA	mode of action
MPPD	multiple path particle dosimetry
MRL	minimum reporting level
Na	sodium
NAFLD	nonalcoholic fatty liver disease
ND	no data
NF-κB	nuclear factor kappa B pathway
NH_4^+	ammonium
NHANES	National Health and Nutrition Examination Survey
NH DES	New Hampshire Department of Environmental Services
NJ DEP	New Jersey Department of Environmental Protection
NMD	normalized mean difference
NOAEL	no-observed-adverse-effect level
NPDWR	National Primary Drinking Water Regulation
NPL	National Priorities List
NR	nuclear receptor
NTP	National Toxicology Program
OCSPP	Office of Chemical Safety and Pollution Prevention
OECD	Organisation for Economic Co-operation and Development
OLEM	Office of Land and Emergency Management
OR	odds ratio

ORD	Office of Research and Development
OSF	oral slope factor
OW	Office of Water
PAC	protective action criteria
PBPK	physiologically based pharmacokinetic
PBTK	physiologically based toxicokinetic
PCL	protective concentration level
PECO	populations, comparators, exposures, and outcomes
PFAS	per- and polyfluoroalkyl substances
PFBA	perfluorobutanoic acid
PFBS	perfluorobutane sulfonate
PFCA	perfluoroalkyl carboxylic acid
PFDA	perfluorodecanoic acid
PFHxA	perfluorohexanoic acid
PFHxS	perfluorohexanesulfonate
PFNA	perfluorononanoic acid
PFOA	perfluorooctanoic acid
PFOS	perfluorooctane sulfonate
PFSA	perfluoroalkane sulfonic acid
PI3K-Akt	phosphatidylinositol-3-kinase-serine/threonine kinase Akt
PK	pharmacokinetic
POD	point of departure
PPARα	peroxisome proliferator-activated receptor alpha
PPRTV	Provisional Peer-Reviewed Toxicity Value
PR	preliminary review
pt.	point
PVDF	polyvinylidene fluoride
PWS	public water system
PXR	pregnane X receptor
RCRA	Resource Conservation and Recovery Act
RfC	inhalation reference concentration
RfD	oral reference dose
ROBINS-I	Risk of Bias in Nonrandomized Studies of Interventions
ROS	reactive oxygen species
RXR	retinoid X receptor
SD	standard deviation
SDWA	Safe Drinking Water Act
$T_{0.5A}$	elimination half-life in animals
$T_{0.5H}$	elimination half-life in humans
TCEQ	Texas Commission on Environmental Quality
TD	toxicodynamic
TDI	tolerable daily intake
TEEL	temporary emergency exposure limit
TNFα	tumor necrosis factor alpha
TRI	Toxics Release Inventory
TSCA	Toxic Substances Control Act
TSCATS	Toxic Substances Control Act Test Submissions
UCMR	Uncontaminated Monitoring Rule
UF	uncertainty factor
Vd	volume of distribution
WHO	World Health Organization
wt.	weight
XME	xenobiotic metabolizing enzymes

1

AUTHORS | CONTRIBUTORS | REVIEWERS

Assessment Teams (U.S. EPA/Center for Public Health and Environmental Assessment [CPHEA])

Michelle Angrish (PFHxA co-lead)

Xabier Arzuaga (PFHxS co-lead)

Thomas Bateson

Allen Davis (PFBA and statistics co-lead)

Laura Dishaw (PFHxA co-lead)

Ingrid L. Druwe (PFHxS co-lead)

J. Phillip Kaiser (PFDA co-lead)

Andrew Kraft (PFNA co-lead)

Lucina Lizarraga (PFDA co-lead)

Pamela Noyes (PFNA co-lead)

Elizabeth Radke (epidemiology lead)

Kristen Rappazzo

Alan Sasso

Paul Schlosser (PBPK and ADME lead)

Michele Taylor (PFBA and PFNA co-lead)

Michael Wright

Jay Zhao (statistics co-lead)

Technical Experts/Contributors (U.S. EPA/CPHEA)

Audrey Galizia

Kelly Garcia (Oak Ridge Associated Universities [ORAU] contractor, no longer with EPA)

Carolyn Gigot (ORAU contractor, no longer with EPA)

Andrew Greenhalgh (ORAU contractor, no longer with EPA)

Belinda Hawkins

Amanda Persad

Linda Phillips (retired)

Brittany Schulz (ORAU contractor)

Andre Weaver

Scott Wesselkamper (no longer with EPA)

Amina Wilkins

George Woodall

Technical Experts/Contributors (outside of U.S. EPA)

John Bucher	National Toxicology Program/Office of Health Assessment and Translation
Andrew Rooney	National Toxicology Program/Office of Health Assessment and Translation/Director
Kyla Taylor	National Toxicology Program/Office of Health Assessment and Translation

Executive Direction

Tina Bahadori	U.S. EPA/Office of Research and Development/Senior Science Advisor
Samantha Jones (former PFAS assessment team lead)	U.S. EPA/CPHEA/Associate Director
Kristina Thayer	U.S. EPA/CPHEA/Chemical and Pollutant Assessment Division (CPAD)/Director
Andrew Kraft (PFAS assessment team lead)	U.S. EPA/CPHEA/CPAD/Senior Science Advisor

Production Team and Review (U.S. EPA/CPHEA)

Anna Chaplin

Madison Feshuk (ORAU contractor)

Catherine Gibbons

Hillary Hollinger (ORAU contractor)

Ryan Jones

Jennifer Nichols

Dahnish Shams

Vicki Soto

1

1. INTRODUCTION

Per- and polyfluoroalkyl substances (PFAS) are a large class of synthetic (man-made) chemicals widely used in consumer products and industrial processes. The basic structure of PFAS consists of a carbon chain surrounded by fluorine atoms, with different chemicals possessing different end groups (see examples in Section 2.1.1); thousands of distinct PFAS exist in commerce. To help address this complex issue, the Environmental Protection Agency (EPA) is taking a proactive approach. Specifically, the development of human health toxicity assessments for exposure to individual PFAS represents only one component of the broader PFAS action plan underway at the U.S. EPA (https://www.epa.gov/pfas/epas-pfas-action-plan). The five toxicity assessments being developed according to the scope and methods outlined in this protocol build upon several other PFAS assessments that have already been developed (see Section 2.1.7).

This protocol document presents the methods for conducting the systematic reviews and dose-response analyses for assessments of perfluorodecanoic acid (PFDA), perfluorononanoic acid (PFNA), perfluorohexanoic acid (PFHxA), perfluorohexanesulfonate (PFHxS), and perfluorobutanoic acid (PFBA), and their related salts (see Figure 1). This includes a summary of why these specific PFAS were prioritized for evaluation, description of the objectives and specific aims of the assessments, draft populations, exposures, comparators, and outcomes (PECO) criteria, and identification of key areas of scientific complexity. This assessment protocol will be posted on the Integrated Risk Information System (IRIS) website (https://cfpub.epa.gov/ncea/iris2/atoz.cfm) for a 45-day comment period. The protocol will also be published in the Zenodo data repository (https://zenodo.org/). Public input received on the protocol is considered during preparation of the draft assessments and any adjustments made to the protocol will be reflected in an updated version released in conjunction with the draft assessments. The literature search results for these five PFAS will also be posted to the Health and Environmental Research Online (HERO) database[1] upon public release of the protocol (the literature search results will be regularly updated during draft development and the subsequent stages of assessment review).

[1]PFBA: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2632
PFHxA: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2628
PFHxS: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2630
PFNA: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2633
PFDA: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2614

2. SCOPING AND PROBLEM FORMULATION SUMMARY

2.1. BACKGROUND

2.1.1. Chemical and Physical Properties

Perfluorodecanoic acid (PFDA; CASRN 335-76-2), perfluorononanoic acid (PFNA; CASRN 375-24-4), perfluorohexanoic acid (PFHxA, CASRN 307-24-4), perfluorohexanesulfonate (PFHxS, CASRN 355-46-4), and perfluorobutanoic acid (PFBA, CASRN 375-22-4), and their related salts, are members of the group PFAS. Section 2.2 ("Scoping Summary") outlines the rationale for why these PFAS were prioritized for assessment. Buck et al. (2011) define PFAS as fluorinated substances that "contain 1 or more C atoms on which all the H substituents (present in the nonfluorinated analogues from which they are notionally derived) have been replaced by F atoms, in such a manner that they contain the perfluoroalkyl moiety ($C_nF_{2n+1}^-$)." More specifically, PFDA, PFNA, PFHxA, and PFBA are classified as perfluoroalkyl carboxylic acids (PFCAs), and PFHxS is a perfluoroalkane sulfonic acid (PFSA) (OECD, 2015). PFCAs containing seven or more perfluorinated carbon groups and PFSAs containing six or more perfluorinated carbon units are considered long-chain PFAS (ATSDR, 2018; OECD, 2015; Buck et al., 2011). Thus, PFDA, PFNA, and PFHxS are considered long-chain PFAS, and PFHxA and PFBA are short-chain PFAS. The chemical structures of PFDA, PFNA, PFHxA, PFHxS, and PFBA, and their related salts, are presented in Figure 1 (along with their CASRNs), and their physiochemical properties are provided in Table 1.

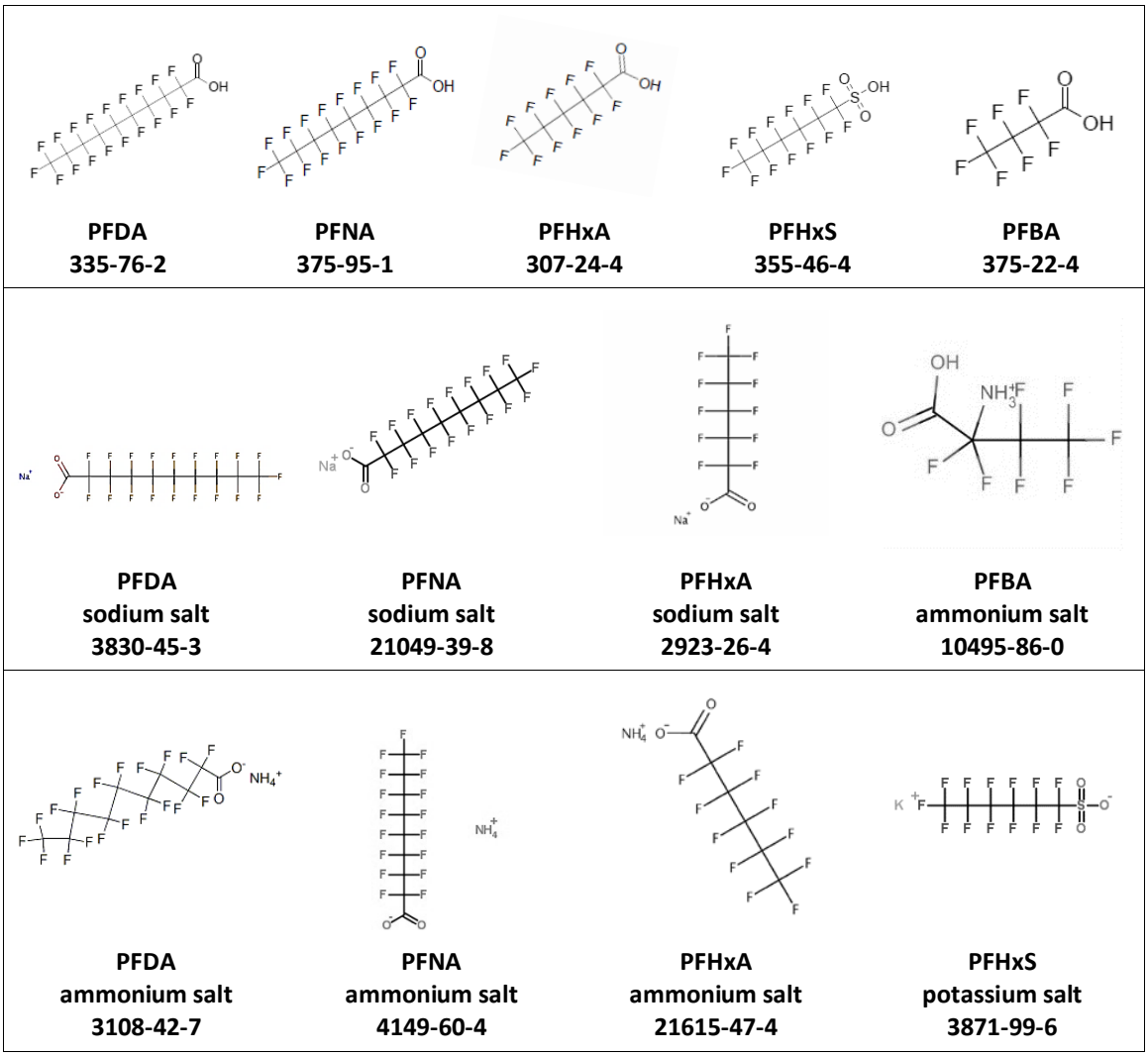


Figure 1. Chemical structures of per- and polyfluoroalkyl substances (PFAS).

1

Table 1. Physiochemical properties of five per- and polyfluoroalkyl substances (PFAS) and their related salts

Property (unit)	PFDA + salts			PFNA + salts			PFHxA	PFHxS + salts		PFBA + salts	
	PFDA[a]	NH_4^+ salt[b]	Na salt[c]	PFNA[d]	NH_4^+ salt[e]	Na salt[f]	PFHxA[g]	PFHxS[h]	K salt[i]	PFBA[j]	NH_4^+ salt[k]
Molecular wt. (g/mol)	514	531	536	464	481	486	314	400	438	214	ND
Melting pt. (°C)	79.5	165*	71.4*	66.5	165*	ND	14.0	190	273	−17.5	ND
Boiling pt. (°C)	218	205*	205*	218	190*	196	157	345	236*	121	ND
Density (g/cm^3)	1.79*	ND	ND	1.78*	ND	ND	1.69*	1.84*	1.84*	1.65	ND
Vapor pressure (mm Hg)	3.89×10^{-2}*	3.51×10^{-2}*	3.51×10^{-2}*	7.06×10^{-2}*	7.38×10^{-2}*	1.71×10^{-1}	9.08×10^{-1}	9.95×10^{-3}*	9.95×10^{-3}*	6.37	ND
Henry's law constant (atm-m^3/mole)	3.55×10^{-10}*	3.55×10^{-10}*	3.55×10^{-10}*	1.64×10^{-10}*	1.64×10^{-10}*	ND	2.51×10^{-10}*	1.96×10^{-10}*	1.96×10^{-10}*	4.99×10^{-5}*	ND
Water solubility (mol/L)	8.3×10^{-6}*	8.41×10^{-1}*	8.41×10^{-1}*	1.38×10^{-5}*	8.45×10^{-1}*	ND	9.39×10^{-5}	6.08×10^{-4}	2.96×10^{-4}*	2.09×10^{-3}	ND
pKa	−0.17*	ND	ND	−0.17*	ND	ND	−0.16	0.14*	ND	0.08*	ND
LogP	7.37*	4.36*	4.36*	6.59*	3.12*	ND	2.51	2.20	2.97*	1.43	ND
Soil adsorption coefficient (L/kg)	397*	397*	397*	2,830*	2,830*	ND	1,070*	2,300*	2,300*	47.9*	ND
Bioconcentration factor	789*	29.8*	29.8*	752*	4.95*	ND	41*	118*	5.94*	7.61*	ND

K = potassium; ND = no data; NH_4^+ = ammonium; Na = sodium; pt. = point; wt. = weight.

[a]CASRN 335-76-2. U.S. EPA (2018a) Chemistry Dashboard (https://comptox.epa.gov/dashboard/; search = PFDA) for all values except pKa (ATSDR, 2018). Values are median experimental values (when available), or median or average predicted values.

[b]CASRN 3108-42-7. Predicted average values from the U.S. EPA (2018a) Chemistry Dashboard (search = 3108-42-7).

[c]CASRN 3830-45-3. Predicted average values from the U.S. EPA (2018a) Chemistry Dashboard (search = 3830-45-3).

[d]CASRN 375-95-1. U.S. EPA (2018a) Chemistry Dashboard (search = PFNA) for all values except pKa (NLM, 2013). Values are median experimental values (when available), or median or average predicted values.

[e]CASRN 4149-60-4. Predicted average values from the U.S. EPA (2018a) Chemistry Dashboard (search = 4149-60-4).

[f]CASRN 21049-39-8. ChemNet website: http://www.chemnet.com/cas/en/21049-39-8/sodium-heptadecafluorononanoate.html for all values.

[g]CASRN: 307-24-4. U.S. EPA (2018a) Chemistry Dashboard (search = 307-24-4) for all values except pKa (NLM, 2016). Values are median or average experimental values (when available), or median or average predicted values.

[h]CASRN 355-46-4. U.S. EPA (2018a) Chemistry Dashboard (search = 355-46-4) for all values except pKa (NLM, 2017). Values are median or average experimental values (when available), or median or average predicted values.

Property (unit)	PFDA + salts			PFNA + salts			PFHxA	PFHxS + salts		PFBA + salts	
	PFDA[a]	NH_4^+ salt[b]	Na salt[c]	PFNA[d]	NH_4^+ salt[e]	Na salt[f]	PFHxA[g]	PFHxS[h]	K salt[i]	PFBA[j]	NH_4^+ salt[k]

[i]CASRN 3871-99-6. U.S. EPA (2018a) Chemistry Dashboard (search = 3871-99-6) for all values. Values are median or average experimental values (when available), or median or average predicted values.
[j]CASRN 375-22-4. U.S. EPA (2018a) Chemistry Dashboard (search = 375-22-4) https://comptox.epa.gov/dashboard/dsstoxdb/results?utf8=%E2%9C%93&search=375-22-4 for all values except pKa (ATSDR, 2018). Values are median or average experimental values (when available), or median or average predicted values.
[k]CASRN 10495-86-0
*Predicted value.

2.1.2. Sources, Production, and Use

PFAS are synthetic (man-made) compounds that have been used since the 1940s in consumer products and industrial applications because of their resistance to heat, oil, stains, grease, and water. They have been used in stain-resistant fabrics for clothing, carpets, and furniture; nonstick cookware; food packaging (e.g., popcorn bags, and fast-food containers); and personal care products (e.g., dental floss, cosmetics, and sunscreen) (ATSDR, 2018). Some PFAS have also been used in firefighting foam and as industrial surfactants, emulsifiers, wetting agents, additives, and coatings, and in the aerospace, automotive, building, and construction industries to help reduce friction (ATSDR, 2018). Because of their widespread use, the release of PFAS into the air, water, and soil, and their persistence in the environment, most people in the United States have been exposed to PFAS (see https://www.epa.gov/pfas/epas-pfas-action-plan for additional details). Examples of how the five PFAS of interest have been used include:

- **PFDA** has been used in stain and grease-proof coatings on food packaging, furniture, upholstery, and carpet (Harbison et al., 2015), and as a lubricant, wetting agent, plasticizer, and corrosion inhibitor (KemI, 2015).

- **PFNA** has been used as a processing aid in the production of fluoropolymers, primarily polyvinylidene fluoride (PVDF), which is a plastic designed to be temperature resistant and chemically nonreactive (NJDWQI, 2017; Prevedouros et al., 2006). It has also been used in aqueous film-forming foam (AFFF) for fire suppression (Laitinen et al., 2014).

- **PFHxA** is not currently a commercial product; it is a breakdown product of "stain- and grease-proof coatings on food packaging and household products" (NTP, 2018b). It has been proposed as a replacement for the commonly used perfluorooctanoic acid (PFOA) and perfluorooctane sulfonate (PFOS) (Klaunig et al., 2015).

- **PFHxS** has been used as a surfactant to make fluoropolymers, and in water- and stain-protective coatings for carpets, paper, and textiles (NTP, 2018a). It may also be present in certain industrial and consumer products such as "fire-fighting foams, food-contact papers, water-proofing agents, cleaning and polishing products either for

intentional uses (as surfactants or surface protection agents) or as unintentional impurities from industrial production processes" (Norwegian Environment Agency, 2018).

- **PFBA** is a breakdown product of other PFAS that are used in stain-resistant fabrics, paper food packaging, and carpets; it was also used for manufacturing photographic film (MDH, 2009).

The U.S. Environmental Protection Agency (EPA) has been working with companies in the fluorochemical industry since the early 2000s to phase out the production and use of long-chain PFAS (https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/risk-management-and-polyfluoroalkyl-substances-pfass). However, the past production and use of these PFAS has resulted in their release to the environment through various waste streams (NLM, 2016, 2013). Also, because products containing PFAS are still in use, they continue to be a source of environmental PFAS contamination through their disposal and subsequent breakdown in the environment (Kim and Kannan, 2007).

Chemical reporting data (CRD) on production volumes are not available in EPA's ChemView (U.S. EPA, 2019) for PFDA, PFNA, PFHxA, PFHxS, PFBA, or their salts. Also, because there are no requirements to report releases to the environment from facilities manufacturing, processing, or otherwise using PFAS, quantitative information is not available in EPA's Toxics Release Inventory (TRI) (U.S. EPA, 2019; ATSDR, 2018).

2.1.3. Environmental Fate and Transport

PFAS are very stable and persistent in the environment (ATSDR, 2018), and many are found worldwide in the environment, wildlife, and humans (https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/risk-management-and-polyfluoroalkyl-substances-pfass). They have been detected at a variety of sites, including private and federal facilities, and have been associated with various sources, including AFFF, chrome-plating facilities, PFAS manufacturers, and industries that use PFAS (e.g., textiles) (ATSDR, 2018). The environmental fate and transport of PFAS potentially can include releases to air to soil and surficial water bodies which can then lead to migration to subsurface soils and ground water contamination (Guelfo et al., 2018)(https://www.atsdr.cdc.gov/pfas/index.html).

PFAS released to air exist in the vapor phase and resist photolysis, but particle-bound concentrations have also been measured (NLM, 2017, 2016, 2013; Kim and Kannan, 2007). The atmospheric half-life for degradation by reaction with photochemically produced hydroxyl radicals is estimated to be 31 days for PFNA and PFHxA, and 115 days for PFHxS (NLM, 2017, 2016, 2013). Long-range atmospheric transport of PFAS is possible, as indicated by the detection of PFHxS in remote arctic and marine air samples (NLM, 2017). Wet and dry deposition are potential removal processes for particle-bound PFAS in air (e.g., to surface water or soil) (ATSDR, 2018). Standardized analytical methods for measuring these five PFAS in ambient air is an area of ongoing research.

In soil, the mobility of PFAS will vary depending on their soil adsorption coefficients (see Table 1), with PFNA predicted to be the least mobile and PFBA the most mobile of the five PFAS addressed here. Volatilization of PFNA, PFHxA, and PFHxS from moist soil is not expected to be an important transport process (NLM, 2017, 2016, 2013). Uptake of soil PFAS to plants can occur (ATSDR, 2018). Yoo et al. (2011) estimated grass-soil accumulation factors (grass concentration divided by soil concentration) of 3.4, 0.12, and 0.10 for PFHxA, PFNA, and PFDA, respectively, based on samples collected from a site with bio-solids-amended soil. Zhao et al. (2016) observed that shorter chain PFAS like PFBA were transported more readily from the roots to the shoots of wheat plants than longer chain PFAS.

PFNA, PFHxA, and PFHxS are expected to adsorb to suspended solids and sediments in water (NLM, 2017, 2016, 2013), and hydrolysis is not expected to be an important fate process (ATSDR, 2018). The potential for PFAS to bioaccumulate in aquatic organisms can be assessed using their bioconcentration factors, with the predicted potential for PFDA and PFNA to bioaccumulate being high compared to PFHxA, PFHxS, and PFBA (see Table 1). As described in Section 2.2, standardized analytical methods for measuring these five PFAS in drinking water exist (for 4 of the 5 PFAS to be assessed) or are under development (i.e., for PFBA). Non-drinking water methods are currently under development.

2.1.4. Environmental Concentrations

PFDA, PFNA, PFHxA, PFHxS, and PFBA have not been evaluated under the National Air Toxics Assessment program (https://www.epa.gov/national-air-toxics-assessment). However, PFDA, PFNA, and PFHxS were measured at concentrations ranging from less than the limit of detection (LOD) to 1.56 pg/m^3 in the vapor and particle phases of air samples collected from an urban area of Albany, NY in 2006 (Kim and Kannan, 2007). PFAS have also been measured in indoor air and dust, and they may be associated with the indoor use of consumer products such as PFAS-treated carpets or other textiles (ATSDR, 2018). For example, Kato et al. (2009) analyzed dust samples collected from 39 homes in the United States, United Kingdom, Germany, and Australia for PFAS, including PFDA, PFNA, PFHxA, and PFHxS. These PFAS were detected in 38.5, 25.6, 46.2, and 79.5% of the samples, respectively. Likewise, Strynar and Lindstrom (2008) analyzed dust samples from 110 homes and 10 day care centers in North Carolina and Ohio, and detected PFDA, PFNA, and PFHxA in 30.4, 42.9, and 92.9% of the samples, respectively. Indoor air samples (N = 4) from a town in Norway had mean concentrations of 3.4 pg/m^3 for PFDA, 2.7 pg/m^3 for PFNA, and <4.1 pg/m^3 for PFHxS (Barber et al., 2007).

The levels of PFAS in soil and sediment surrounding perfluorochemical industrial facilities have been measured at concentrations ranging from less than the LOD to 124 ng/g for PFBA and less than the LOD to 3,470 ng/g for PFHxS (ATSDR, 2018). PFDA, PFNA, PFHxA, PFHxS, and PFBA were also detected at an Australian training ground where AFFFs had been used (Baduel et al., 2015). PFDA, PFNA, PFHxA, PFHxS, and PFBA were detected at 10 U.S. military sites in 67.0, 71.4, 70.3, 76.9, and 38.5% of the surface soil samples, respectively, and 48.5, 12.1, 63.6, 72.7, and 24.2%

of the sediment samples, respectively (ATSDR, 2018). Table 2 shows the concentrations of these PFAS in soil and sediment at these military sites.

EPA conducted monitoring for several PFAS in drinking water as part of the third Uncontaminated Monitoring Rule (UCMR) (U.S. EPA, 2016e). Under the UCMR, all public water systems (PWSs) serving more than 10,000 people and a representative sample of 800 PWSs serving 10,000 or fewer people were monitored for 30 unregulated contaminants between January 2013 and December 2015. PFNA and PFHxS were among the 30 contaminants monitored. PFNA was detected above the minimum reporting level (MRL) of 0.02 µg/L in 14 of the 4,920 PWSs tested and in 19 of the 36,972 samples collected. PFHxS was detected above the MRL of 0.03 µg/L in 55 of the 4,920 PWSs tested and in 207 of the 36,971 samples collected. UCMR data were not available for PFDA, PFHxA, or PFBA. However, samples from seven municipal wells in Oakdale, MN were analyzed for PFHxA and PFBA. The concentrations ranged from <0.025 to 0.235 µg/L and 0.0855 to 2.04 µg/L, respectively (U.S. EPA, 2017b). Kim and Kannan (2007) analyzed lake water, rain water, snow, and surface water from Albany, NY, and reported concentrations of PFDA, PFNA, and PFHxS ranging from less than the LOD to 0.0135 µg/L. PFAS were detected at higher concentrations in groundwater samples from an industrial site (3M Cottage Grove) in Minnesota. PFHxS and PFBA were detected in all 7 wells that were sampled at concentrations ranging from 6.47–40 µg/L and 23.3–318 µg/L, respectively [(WS, 2007) as cited in (ATSDR, 2018)]. The concentrations of these five PFAS measured at National Priorities List (NPL) sites are shown in Table 3, and the concentrations of PFAS measured in surface water and groundwater at 10 military installations are given in Table 2.

Table 2. Per- and polyfluoroalkyl substances (PFAS) levels at 10 military installations

Media	Measure	PFDA	PFNA	PFHxA	PFHxS	PFBA
Surface soil	Frequency of detection (%)	67.03	71.43	70.33	76.92	38.46
	Reporting limit (µg/kg)	0.28	0.23	0.16	0.29	0.12
	Median (µg/kg)	0.980	1.30	1.75	5.70	1.00
	Maximum (µg/kg)	15.0	23.0	51.0	1,300	31.0
Subsurface soil	Frequency of detection (%)	12.50	14.42	65.38	59.62	29.81
	Reporting limit (µg/kg)	0.30	0.24	0.16	0.31	0.13
	Median (µg/kg)	1.40	1.50	1.04	4.40	0.960
	Maximum (µg/kg)	9.40	6.49	140	520	14.0
Sediment	Frequency of detection (%)	48.48	12.12	63.64	72.73	24.24
	Reporting limit (µg/kg)	0.46	0.38	0.26	0.48	0.21
	Median (µg/kg)	1.90	1.10	1.70	9.10	1.70
	Maximum (µg/kg)	59.0	59.0	710	2,700	140
Surface water	Frequency of detection (%)	52.00	36.00	96.00	88.00	84.00
	Reporting limit (µg/L)	0.008	0.017	0.003	0.007	0.010
	Median (µg/L)	0.067	0.096	0.320	0.710	0.076
	Maximum (µg/L)	3.20	10.0	292	815	110
Groundwater	Frequency of detection (%)	34.78	46.38	94.20	94.93	85.51
	Reporting limit (µg/L)	0.008	0.018	0.003	0.007	0.010
	Median (µg/L)	0.023	0.105	0.820	0.870	0.180
	Maximum (µg/L)	1.80	3.00	120	290	64.0

Source: (ATSDR, 2018).

1

Table 3. Per- and polyfluoroalkyl substances (PFAS) levels in water, soil, and air at National Priorities List sites

Media	Measure	PFDA	PFNA	PFHxA	PFHxS	PFBA
Water	Median (ppb)	ND	ND	0.25	0.26	2.15
	Geometric mean (ppb)	ND	ND	0.10	1.12	1.03
Soil	Median (ppb)	ND	27.2	1,175	5,585	1,600
	Geometric mean (ppb)	ND	27.2	1,175	5,585	1,600
Air	Median (ppbv)	ND	ND	ND	ND	ND
	Geometric mean (ppbv)	ND	ND	ND	ND	ND

ND = no data.
Source: (ATSDR, 2018).

2

3 Schecter et al. (2012) collected 10 samples of 31 food items from five grocery stores in
4 Texas and analyzed them for persistent organic pollutants, including PFDA, PFNA, PFHxA, and
5 PFHxS. PFDA, PFNA, and PFHxA were not detected in any of the foods targeted, and PFHxS was
6 detected in cod fish at a concentration of 0.07 ng/g wet weight. PFAS have been detected in fish

from U.S. lakes and rivers with concentrations ranging from less than the limit of quantification to 15.0 ng/g for PFDA, and <1 to 0.47 ng/g for PFHxS (ATSDR, 2018). Stahl et al. (2014) characterized PFAS in freshwater fish from 164 U.S. urban river sites and 157 Great Lakes sites. PFDA, PFNA, PFHxA, PFHxS, and PFBA were detected in 92, 69, 15, 45, and 16% of the samples, at maximum concentrations of 13.0, 9.7, 0.8, 3.5, and 1.3 ng/g, respectively. Apart from fish, overall dietary data for the United States are limited; however, Schaider et al. (2017) detected PFASs in food packaging collected from U.S. fast food restaurants. Data from other countries (e.g., South Korea, Brazil, Saudi Arabia) suggest that these PFAS can sometimes be detected in samples of food products including shellfish, dairy products, meats, vegetables, food packaging materials, and water (both tap and bottled) (Chen et al., 2018; Surma et al., 2017; Heo et al., 2014; Moreta and Tena, 2014; Pérez et al., 2014). The relevance of these detects (and the associated PFAS levels) to U.S. products is unknown.

2.1.5. Potential for Human Exposure

The general population may be exposed to PFAS through multiple routes, including ingestion of drinking water and food, ingestion of dust, hand-to-mouth and dermal transfer in products and materials containing these chemicals, and inhalation via indoor and outdoor air (ATSDR, 2018; NLM, 2017, 2013). The oral route of exposure has been considered the most important route of exposure among the general population for PFAS (Klaunig et al., 2015).

The presence of PFAS in human blood provides evidence of exposure among the general population. PFAS have been monitored in the human population as part of the National Health and Nutrition Examination Survey (NHANES). PFDA, PFNA, and PFHxS were measured in serum samples collected in 2013–2014 from more than 2,000 survey participants. The results of these analyses are presented in Table 4. PFDA and PFNA have also been observed in cord blood and human milk (ATSDR, 2018). Pinney et al. (2014) and Papadopoulou et al. (2016) observed associations between breastfeeding and elevated levels of PFHxS in the blood of children.

Table 4. Serum per- and polyfluoroalkyl substances (PFAS) concentrations based on National Health and Nutrition Examination Survey (NHANES) 2013–2014 data (µg/L)

Population group	Measure	PFDA	PFNA	PFHxA	PFHxS	PFBA
Total population (*N* = 2,168)	Geometric mean	0.185	0.675	ND	1.35	ND
	50th percentile	0.200	0.700	ND	1.40	ND
	95th percentile	0.700	2.00	ND	5.60	ND
3 to 5 yr (*N* = 181)	Geometric mean	-[a]	0.764	ND	0.715	ND
	50th percentile	0.100	0.620	ND	0.740	ND
	95th percentile	0.370	3.49	ND	1.62	ND
6 to 11 yr (*N* = 458)	Geometric mean	-[a]	0.809	ND	0.913	ND
	50th percentile	<LOD	0.750	ND	0.850	ND
	95th percentile	0.350	3.19	ND	4.14	ND

12 to 19 yr (*N* = 402)	Geometric mean 50th percentile 95th percentile	0.136 0.100 0.400	0.599 0.500 2.00	ND ND ND	1.27 1.10 6.30	ND ND ND
20 yr and older (*N* = 1,766)	Geometric mean 50th percentile 95th percentile	0.193 0.200 0.800	0.685 0.700 2.00	ND ND ND	1.36 1.40 5.50	ND ND ND

LOD = limit of detection; 0.1 (µg/L); ND = no data.
[a]Not calculated because the proportion of results below the limit of detection was considered too high to provide a valid result.
Source: CDC (2018). Fourth National Report on Human Exposure to Environmental Chemicals.

2.1.6. Populations and Lifestages with Potentially Greater Exposures

Populations and lifestages that may experience exposures greater than those of the general population include individuals in occupations that require frequent contact with PFAS-containing products, such as firefighters or individuals who install and treat carpets (ATSDR, 2018), as well as infants and young children (due to their increased hand-to-mouth behaviors). Rotander et al. (2015) analyzed serum samples from 149 Australian firefighters at an AFFF training facility. Mean and median PFHxS concentrations were 10 to 15 times higher than those of the general population of Australia and Canada. Laitinen et al. (2014) evaluated eight firefighters' exposure to PFAS after three training sessions in Finland in which AFFF had been used. The authors found that the firefighters' "serum PFHxS and PFNA concentrations seemed to increase during the three training sessions although they were not the main PFAS in used AFFF." Populations living near fluorochemical facilities where environmental contamination has occurred may also be more highly exposed (ATSDR, 2018). Also, because PFDA can be found in ski wax, individuals who engage in professional ski waxing may be more highly exposed because PFAS in dust may become airborne and inhaled during this process (Harbison et al., 2015).

Populations that rely primarily on seafood for most of their diet, possibly including some native American tribes (Byrne et al., 2017), may also be disproportionately exposed. Christensen et al. (2017) and Haug et al. (2010) used data on serum PFAS levels and 30-day self-reported fish and shellfish ingestion rates from NHANES 2007–2014 to explore potential relationships between PFAS exposures and fish consumption. PFDA, PFNA, and PFHxS were among the PFAS detected in the serum of at least 30% of the NHANES participants, and after adjusting for demographic characteristics, total fish consumption was associated with elevated serum PFDE and PFNA. Shellfish consumption was associated with elevations of all the PFASs examined.

2.1.7. Other Environmental Protection Agency (EPA) Assessments of Per- and Polyfluoroalkyl Substances (PFAS)

EPA released two PFAS assessments for peer review in 2018. Specifically, the draft assessments of (1) 2,3,3,3-tetrafluoro-2-(1,1,2,2,3,3,3-heptafluoropropoxy)propanoic acid (also called hexafluoropropylene oxide [HFPO] dimer acid) (CASRN 13252-13-6) and its ammonium salt

2,3,3,3-tetrafluoro-2-(1,1,2,2,3,3,3-heptafluoropropoxy)propanoate (also called HFPO dimer acid) (CASRN 62037-80-3) referred to as GenX chemicals and (2) perfluorobutane sulfonic acid (CASRN 375-73-5) and its potassium salt potassium perfluorobutane sulfonate (CASRN 29420-49-3) referred to as PFBS. These assessments summarize the available data on the potential human health effects of lifetime exposure to these PFAS and included oral reference doses (RfDs), which estimate (with uncertainty spanning perhaps an order of magnitude) a level of daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious noncancer health effects during a lifetime, and qualitative descriptions of the carcinogenic potential of the chemicals. The PFBS assessment updates a Provisional Peer-Reviewed Toxicity Value (PPRTV) assessment that was developed in support of the Superfund Program and published in 2014 (PFBS PPRTV 2014). In addition, EPA released Drinking Water Health Advisories for PFOA and PFOS in 2016, along with health effect support documents (Drinking Water Health Advisories for PFOA and PFOS). Health advisories are non-enforceable and non-regulatory summaries of technical information on contaminants that can cause human health effects and are known or anticipated to occur in drinking water.

2.2. SCOPING SUMMARY

Given the numerous PFAS of potential interest to the Agency, an extensive scoping effort was undertaken to prioritize PFAS for review. This effort was coordinated across EPA program and regional offices, where staff discussed specific assessment needs as well as the timeliness of those needs. While additional factors were considered during this scoping effort, Table 5 summarizes the primary considerations for selecting the five PFAS described in this protocol, as well as two other PFAS that were recently assessed by EPA: PFBS and GenX chemicals (https://www.epa.gov/pfas/genx-and-pfbs-draft-toxicity-assessments). In short, these PFAS:

- were identified as a priority to inform decision making for EPA's Office of Water (OW), Office of Land and Emergency Management (OLEM), Office of Chemical Safety and Pollution Prevention (OCSPP), Office of Children's Health Protection (OCHP), EPA's regional offices, tribes, or state departments of environmental protection. Most of these PFAS were a priority for multiple patrons;
- had been evaluated in *in vivo* studies of animals and thus might be used to derive toxicity values; and
- had existing (or under development) standardized analytical methods to monitor environmental levels to allow for site-specific application of toxicity values to regulatory decision making.

Table 5. Environmental Protection Agency (EPA) considerations for the selection of per- and polyfluoroalkyl substances (PFAS) for evaluation

PFAS	EPA interest	Animal dose-response data available[a]	Analytical detection methods available[b]	
			Standards	Methods
PFDA	• OLEM priority[c]	Yes	Yes	Yes
PFNA	• OLEM priority[c] • OW (UCMR) priority • Found in industrial effluent and AFFF	Yes	Yes	Yes
PFHxA	• OCSPP priority[d] • OLEM priority[c] • Region 4 (Coosa and Tennessee Rivers) • Found in AFFF	Yes	Yes	Yes
PFHxS	• OCSPP priority • OLEM priority[c] • OW (UCMR) priority • Region 4 (Tennessee River) • Found in AFFF	Yes	Yes	Yes
PFBA	• OLEM priority[c] • OCSPP priority[d] • Found in AFFF	Yes	Yes	Under development
PFBS	• OLEM priority[c] • OCSPP priority[d] • OW (UCMR) priority • Found in AFFF	Yes	Yes	Yes
GenX chemicals	• OCSPP priority[e] • Region 3 priority • Region 4 priority	Yes	Yes	Yes

GenX= perfluoro(2-methyl-3-oxahexanoic) acid (CASRN 13252-13-6); Unknown = status of validated standards and methods was unknown at scoping.

[a]A survey of publicly available literature on PFAS other than PFOA and PFOS (i.e., a broad PubMed search and review of recent assessments, including ATSDR (2018) was performed to identify *in vivo* animal studies that

PFAS	EPA interest	Animal dose-response data available[a]	Analytical detection methods available[b]	
			Standards	Methods

tested multiple PFAS exposure levels and evaluated health endpoints. The quality of the studies was not evaluated, and while multiple PFAS are evaluated in human studies, this was not a focus of the survey.

[b]As of March 2019. The methods noted are for drinking water; non-drinking water methods are being developed.

[c]Found at sites, including private and federal facilities and from various sources, including AFFF, chrome-plating facilities, PFAS manufacturers, and industries that use PFAS (e.g., textiles and electronics). These PFAS have also been detected in environmental media (e.g., surface water; biota).

[d]A significant number of new chemicals submitted to EPA are based on C6 and C4 chemistry. OCSPP often evaluates risk for these compounds based on PFHxA and PFBS, which are the terminal degradation products of certain C6 and C4 compounds.

[e]Replacement for PFOA (e.g., for emulsifiers) and perfluoroethers. GenX chemicals are of concern based on occurrence in NC and because EPA has received requests to review similar types of compounds (e.g., longer chain ethers that might break down to GenX chemicals) as new chemicals.

As described in Section 2.1.5, exposure to these five PFAS can occur via the oral, inhalation, and dermal routes, with oral (e.g., through diet and drinking water) being the predominant one (Klaunig et al., 2015). Given the potential regulatory applications of these PFAS assessments (see Table 6), these assessments will consider PFAS exposures from all exposure routes. The assessments will consider all potential health effects of exposure, both cancer and noncancer.

Table 6. Potential Environmental Protection Agency (EPA) needs and applications for five per- and polyfluoroalkyl substances (PFAS)

EPA program or regional office	PFAS[a]	Oral	Inhalation	Dermal	Potential regulatory application and explanation (at the time scoping was conducted)
OLEM (in coordination with EPA Regions 1–10)	PFDA PFNA PFHxA PFHxS PFBA	✓	✓	✓	Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA). CERCLA authorizes EPA to conduct short- or long-term cleanups at Superfund sites and later recover cleanup costs from potentially responsible parties under Section 107. PFAS toxicological information may be used to make risk determinations for response actions (e.g., short-term removals, long-term remedial response actions). An evaluation of potential actions at Superfund sites considers all routes of exposure. Resource Conservation and Recovery Act (RCRA). RCRA can be drawn upon to help address waste management and cleanup needs, including accidental releases from potentially hazardous waste management facilities.
OW	PFNA PFHxS	✓			Safe Drinking Water Act (SDWA) and Clean Water Act (CWA). The SDWA requires EPA to periodically review the National Primary Drinking Water Regulation (NPDWR) for

EPA program or regional office	PFAS[a]	Oral	Inhalation	Dermal	Potential regulatory application and explanation (at the time scoping was conducted)
					each contaminant and revise the regulation, if appropriate. These potential applications focus on oral exposure.
OCSPP	PFHxA PFHxS PFBA	✓	✓		New chemical submissions to the Office of Pollution Prevention and Toxics within OCSPP.
Region 4	PFHxA PFHxS	✓			Resource Conservation and Recovery Act (RCRA). RCRA can be drawn upon to help address waste management and cleanup needs, including accidental releases from potentially hazardous waste management facilities. For PFAS, the primary concern is potential oral exposure from rivers in Region 4.
OCHP	PFDA PFNA PFHxA PFHxS PFBA	✓	✓	✓	Executive Order 13045—Protection of Children from Environmental Health Risks and Safety Risks: Policy on Evaluating Health Risks to Children. In accordance with EPA's 1995 policy and EO 13045, EPA instituted and reaffirmed an Agency-wide commitment to "consider the risks to infants and children consistently and explicitly as part of risk assessments generated during its decision-making process."

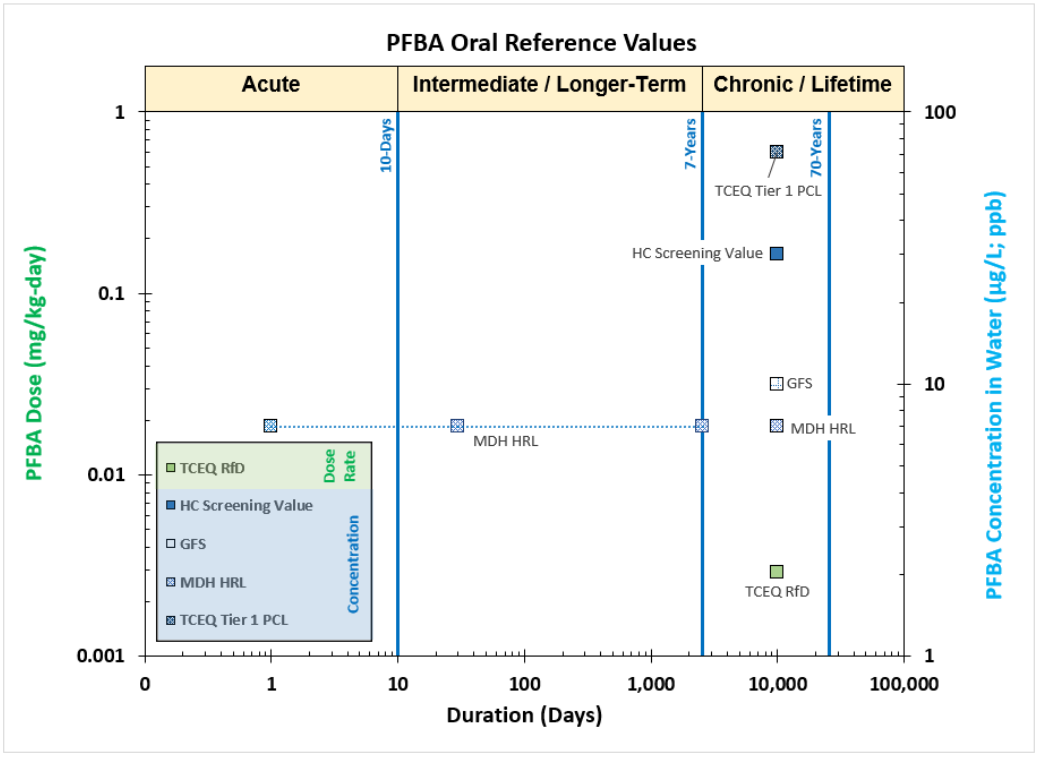
[a]PFAS to which this protocol applies (i.e., excluding PFBS and GenX chemicals).

2.3. PROBLEM FORMULATION

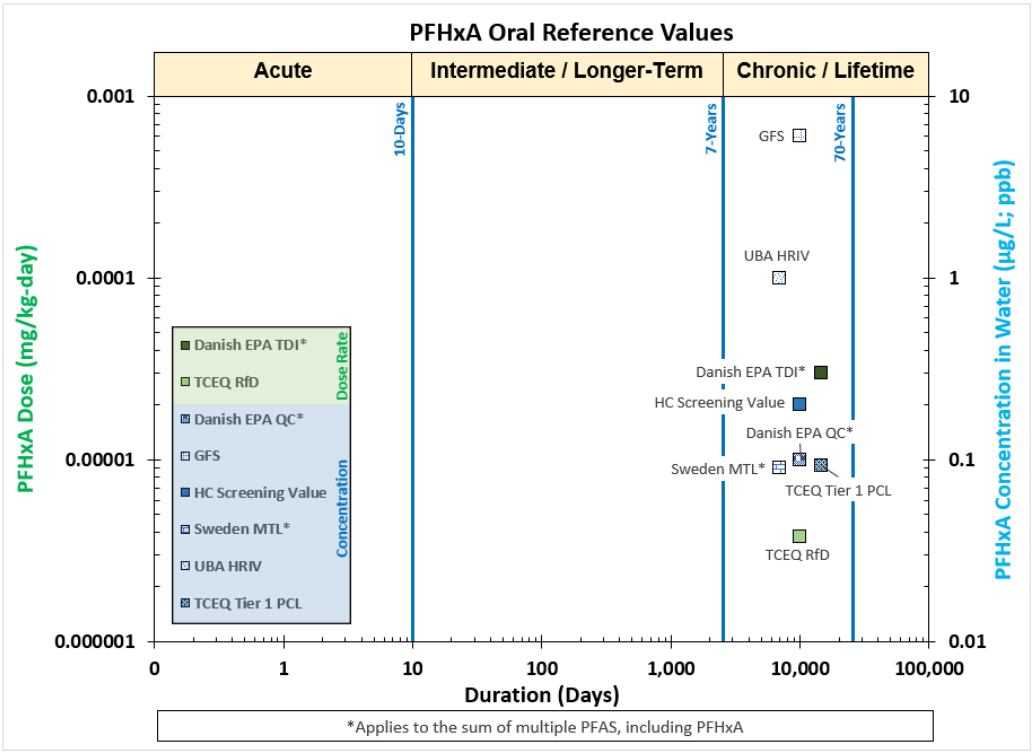
2.3.1. Assessments and Toxicity Values from Other Sources

For the five PFAS addressed in this protocol, a summary of existing values from national, international, and state agencies, as well as U.S. state action levels (current as of March 2019), is provided in Figure 2. It is important to note that these values are not all directly comparable, as some values take into account the potential for human exposure or other considerations. The majority of current values are noncancer toxicity values based on oral exposure studies in rodents, although a few inhalation toxicity values exist (see Table A-1 in Appendix A for more details).

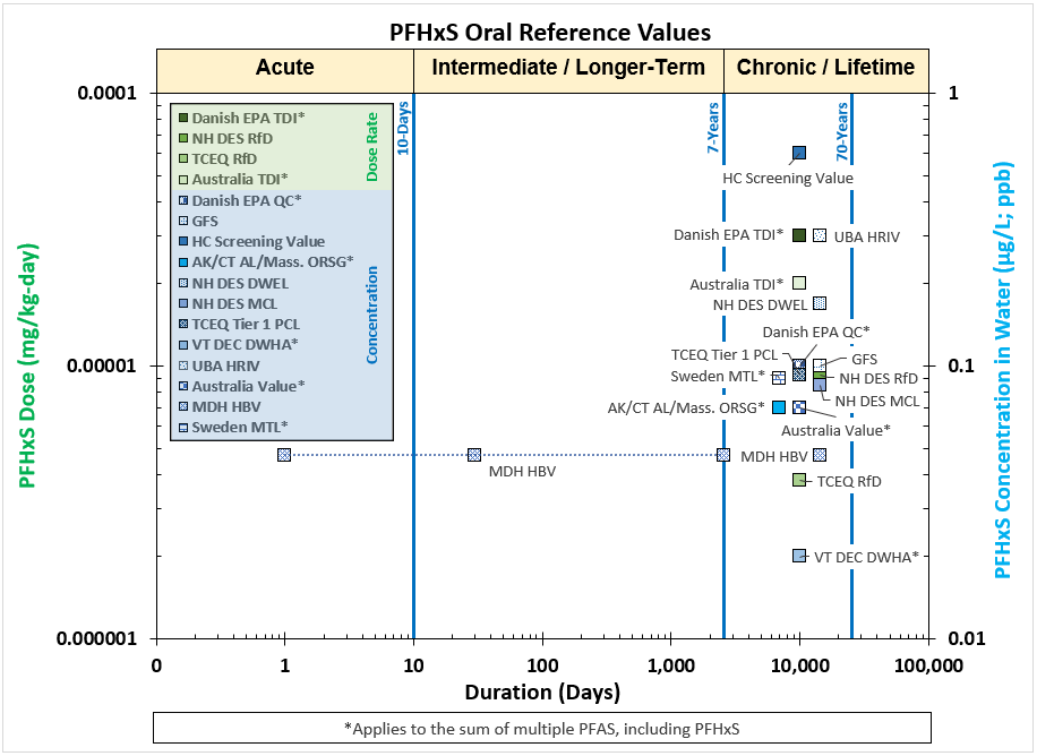
(a)


PFBA Oral Reference Values
Acute
Intermediate / Longer-Term
Chronic / Lifetime
10-Days
7-Years
70-Years
TCEQ Tier 1 PCL
HC Screening Value
GFS
MDH HRL
MDH HRL
TCEQ RfD
TCEQ RfD
HC Screening Value
GFS
MDH HRL
TCEQ Tier 1 PCL
Dose Rate
Concentration
PFBA Dose (mg/kg-day)
PFBA Concentration in Water (µg/L; ppb)
1
0.1
0.01
0.001
100
10
1
0
1
10
100
1,000
10,000
100,000
Duration (Days)

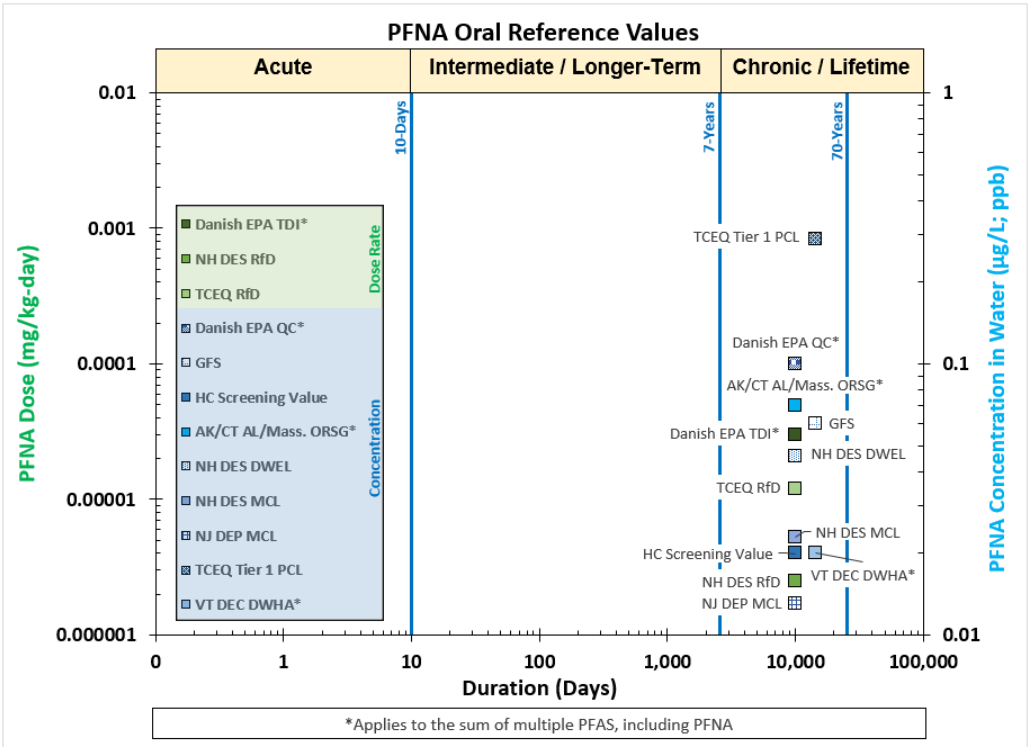
(b)


PFHxA Oral Reference Values
Acute
Intermediate / Longer-Term
Chronic / Lifetime
10-Days
7-Years
70-Years
GFS
UBA HRIV
Danish EPA TDI*
HC Screening Value
Danish EPA QC*
Sweden MTL*
TCEQ Tier 1 PCL
TCEQ RfD
Danish EPA TDI*
TCEQ RfD
Danish EPA QC*
GFS
HC Screening Value
Sweden MTL*
UBA HRIV
TCEQ Tier 1 PCL
Dose Rate
Concentration
PFHxA Dose (mg/kg-day)
PFHxA Concentration in Water (µg/L; ppb)
0.001
0.0001
0.00001
0.000001
10
1
0.1
0.01
0
1
10
100
1,000
10,000
100,000
Duration (Days)
*Applies to the sum of multiple PFAS, including PFHxA

(c)


PFHxS Oral Reference Values
Acute
Intermediate / Longer-Term
Chronic / Lifetime
PFHxS Dose (mg/kg-day)
PFHxS Concentration in Water (µg/L; ppb)
0.0001
0.00001
0.000001
1
0.1
0.01
10-Days
7-Years
70-Years
Dose Rate
Concentration
Danish EPA TDI*
NH DES RfD
TCEQ RfD
Australia TDI*
Danish EPA QC*
GFS
HC Screening Value
AK/CT AL/Mass. ORSG*
NH DES DWEL
NH DES MCL
TCEQ Tier 1 PCL
VT DEC DWHA*
UBA HRIV
Australia Value*
MDH HBV
Sweden MTL*
0
1
10
100
1,000
10,000
100,000
Duration (Days)
*Applies to the sum of multiple PFAS, including PFHxS

(d)


PFNA Oral Reference Values
Acute
Intermediate / Longer-Term
Chronic / Lifetime
PFNA Dose (mg/kg-day)
PFNA Concentration in Water (µg/L; ppb)
0.01
0.001
0.0001
0.00001
0.000001
1
0.1
0.01
10-Days
7-Years
70-Years
Dose Rate
Concentration
Danish EPA TDI*
NH DES RfD
TCEQ RfD
Danish EPA QC*
GFS
HC Screening Value
AK/CT AL/Mass. ORSG*
NH DES DWEL
NH DES MCL
NJ DEP MCL
TCEQ Tier 1 PCL
VT DEC DWHA*
0
1
10
100
1,000
10,000
100,000
Duration (Days)
*Applies to the sum of multiple PFAS, including PFNA

(e)

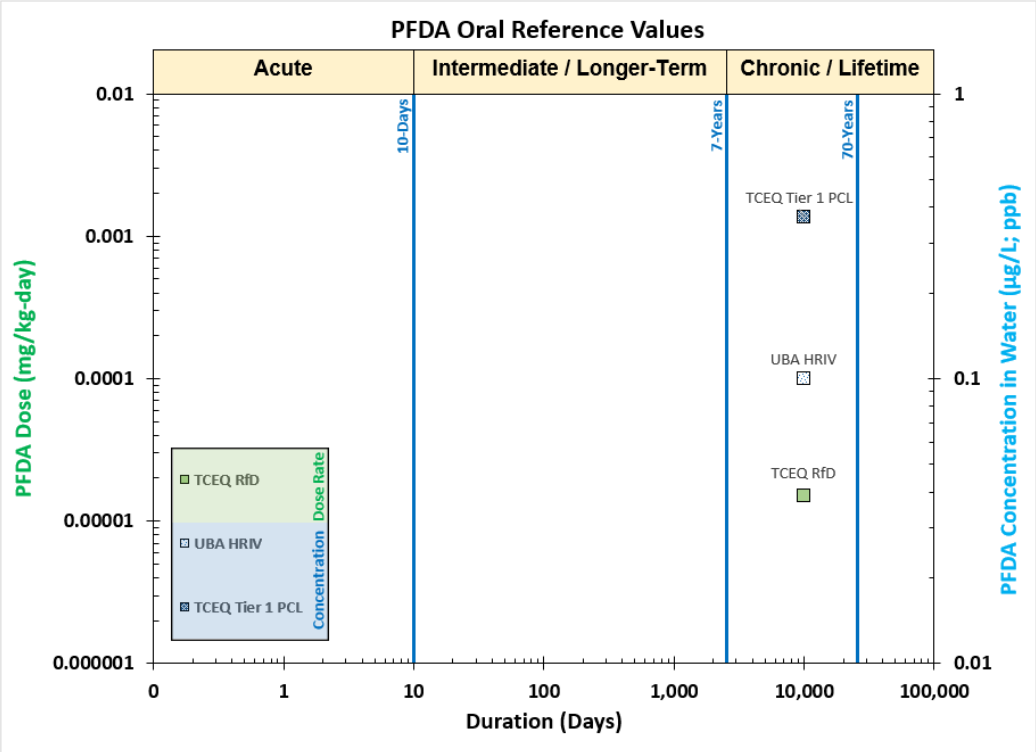


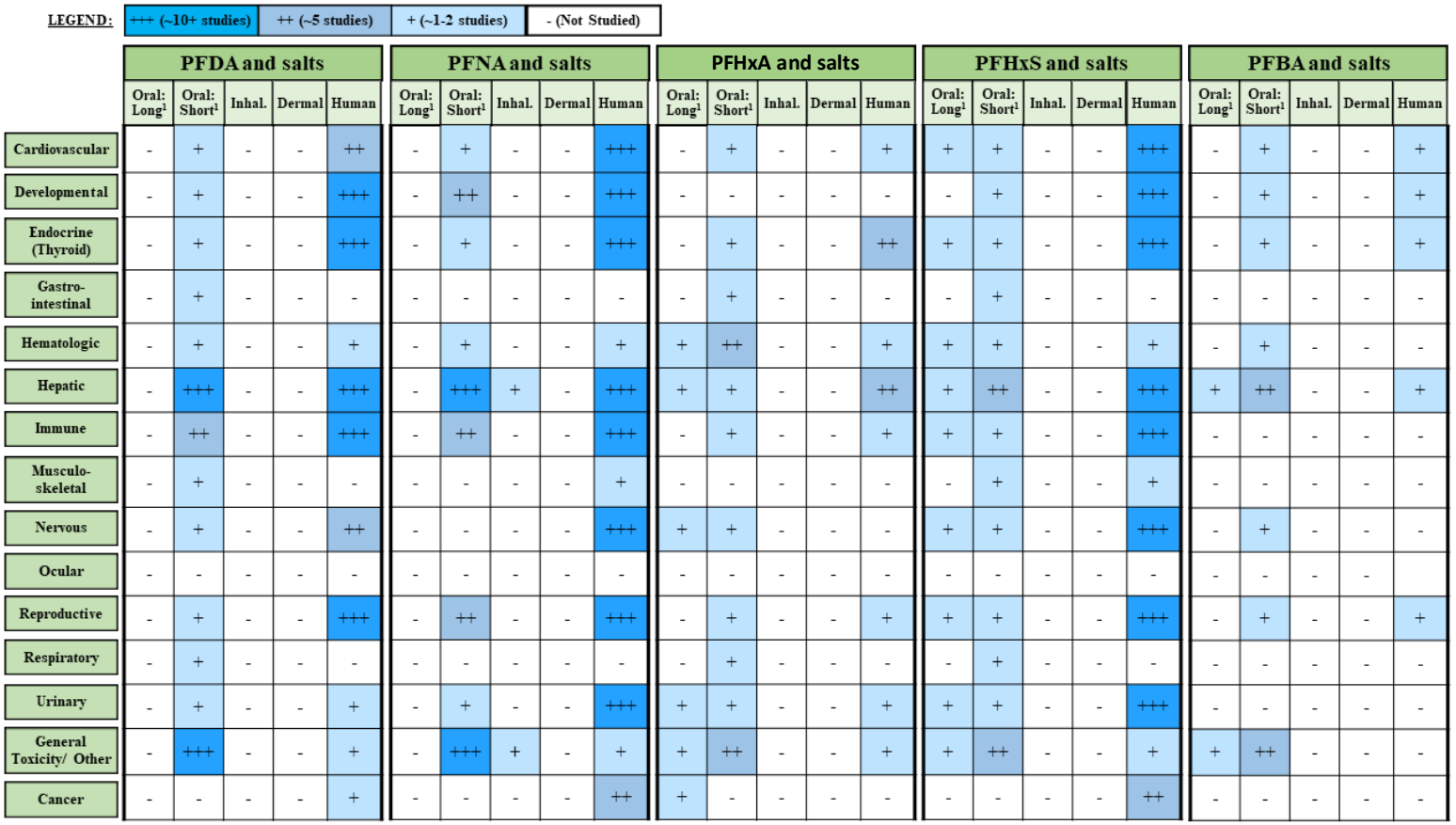
Figure 2. Existing Oral Reference Values for (a) perfluorobutanoic acid (PFBA), (b) perfluorohexanoic acid (PFHxA), (c) perfluorohexanesulfonate (PFHxS), (d) perfluorononanoic acid (PFNA), and (e) perfluorodecanoic acid (PFDA).

2.3.2. Preliminary Literature Inventory for the Five Per- and Polyfluoroalkyl Substances (PFAS) Being Assessed

As described in Section 2.1.1, several of these five PFAS have associated salts of potential interest for human health assessment. Thus, the assessments will address each PFAS as follows:

- **PFBA**: PFBA (CASRN 375-22-4); PFBA ammonium salt (CASRN 10495-86-0)
- **PFHxA**: PFHxA (CASRN 307-24-4); PFHxA ammonium salt (CASRN 21615-47-4); PFHxA sodium salt (CASRN 2923-26-4)
- **PFHxS**: PFHxS (CASRN 355-46-4); PFHxS potassium salt (CASRN 3871-99-6)
- **PFNA**: PFNA (CASRN 375-95-1); PFNA ammonium salt (CASRN 4149-60-4); PFNA sodium salt (CASRN 21049-39-8)
- **PFDA**: PFDA (CASRN 335-76-2); PFDA ammonia salt (CASRN 3108-42-7); PFDA sodium salt (CASRN 3830-45-3)

The results of a preliminary literature inventory of health effect-related studies on these five PFAS and their associated salts are presented in Figure 3. The studies summarized in this preliminary literature inventory are described on the project pages for these assessments in HERO (https://hero.epa.gov; see Section 1 for links to the specific Health and Environmental Research Online [HERO] pages).



LEGEND: +++ (~10+ studies) | ++ (~5 studies) | + (~1-2 studies) | - (Not Studied)

	PFDA and salts					PFNA and salts					PFHxA and salts					PFHxS and salts					PFBA and salts				
	Oral: Long[1]	Oral: Short[1]	Inhal.	Dermal	Human	Oral: Long[1]	Oral: Short[1]	Inhal.	Dermal	Human	Oral: Long[1]	Oral: Short[1]	Inhal.	Dermal	Human	Oral: Long[1]	Oral: Short[1]	Inhal.	Dermal	Human	Oral: Long[1]	Oral: Short[1]	Inhal.	Dermal	Human
Cardiovascular	-	+	-	-	++	-	+	-	-	+++	-	+	-	-	+	+	+	-	-	+++	-	+	-	-	+
Developmental	-	+	-	-	+++	-	++	-	-	+++	-	-	-	-	-	-	+	-	-	+++	-	+	-	-	+
Endocrine (Thyroid)	-	+	-	-	+++	-	+	-	-	+++	-	+	-	-	++	+	+	-	-	+++	-	+	-	-	+
Gastro-intestinal	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-
Hematologic	-	+	-	-	+	-	+	-	-	+	+	++	-	-	+	+	+	-	-	+	-	+	-	-	-
Hepatic	-	+++	-	-	+++	-	+++	+	-	+++	+	+	-	-	++	+	++	-	-	+++	+	++	-	-	+
Immune	-	++	-	-	+++	-	++	-	-	+++	-	+	-	-	+	+	+	-	-	+++	-	-	-	-	-
Musculo-skeletal	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-
Nervous	-	+	-	-	++	-	-	-	-	+++	+	+	-	-	-	+	+	-	-	+++	-	+	-	-	-
Ocular	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Reproductive	-	+	-	-	+++	-	++	-	-	+++	-	+	-	-	+	+	+	-	-	+++	-	+	-	-	+
Respiratory	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-
Urinary	-	+	-	-	+	-	+	-	-	+++	+	+	-	-	+	+	+	-	-	+++	-	-	-	-	-
General Toxicity/ Other	-	+++	-	-	+	-	+++	+	-	+	+	++	-	-	+	+	++	-	-	+	+	++	-	-	-
Cancer	-	-	-	-	+	-	-	-	-	++	+	-	-	-	-	-	-	-	-	++	-	-	-	-	-

Figure 3. Results of a preliminary literature inventory of five per- and polyfluoroalkyl substances (PFAS). Data are approximated based on a cursory review of the literature search results for studies published through 2018 (see Section 4 for details). Health effects are based on groupings from EPA Integrated Risk Information System (IRIS) website (https://cfpub.epa.gov/ncea/iris/search/index.cfm).[a] For this summary, metabolic effects are captured under "other" and "hepatic" includes lipid and lipoprotein measures.

[a]"Oral: long" indicates subchronic or chronic oral exposure duration studies in animals and "Oral: short" reflects short-term and acute oral exposure studies in animals, as well as reproductive and developmental studies.

Based on the results from the preliminary literature inventory in Figure 3, the following health effects appear to be well studied for most PFAS of interest:

- Developmental effects
- Endocrine (primarily thyroid hormone) effects
- Hepatic effects, including lipid and lipoprotein measures
- Immune effects
- Reproductive effects in males or females
- Urinary effects
- General toxicity

As also shown in Figure 3, no studies of dermal exposure were identified. In addition, data are sparse for assessing the potential health effects from chronic or subchronic oral exposure or for inhalation exposure of any duration. Few studies have examined whether exposure to these PFAS may result in carcinogenicity.

Given the potential future utility of comparing evidence across PFAS assessments (including their respective data gaps), the five PFAS assessments will specifically address each of the potential health effects enumerated above as "well studied." In addition, the potential for carcinogenicity will be explicitly addressed in each assessment. Data on several other, variably studied endpoints (i.e., cardiovascular effects, hematological effects, metabolic effects including diabetes, and nervous system effects) will also be summarized when available. These summaries may be developed in association with one of the health effects noted above, as a separate formal evaluation of hazard, or as part of a qualitative summary on "other effects," depending on the assessment-specific data. Given the paucity of available studies and the absence of exceptional evidence in the available studies, information on other health effects (i.e., gastrointestinal effects; musculoskeletal effects; ocular effects; and respiratory effects) may be briefly summarized but will not be formally evaluated in any of these assessments. New literature relating to these outcomes will be monitored during literature search updates for potential inclusion.

2.4. KEY SCIENCE ISSUES

This section describes critical areas of scientific complexity that were identified based on the preliminary literature inventory results summarized in the previous section. These scientific issues are essential to consider during development of these assessments, and the specific methods for doing so within these PFAS assessments are described in subsequent sections.

2.4.1. Toxicokinetic Differences across Species and Sexes

In humans, PFAS generally tend to remain unchanged in the body for long durations (in general, while PFAS bind to macromolecules such as albumin and lipoproteins, they do not undergo internal chemical reactions, and many are not metabolized). They are typically not stored in body fat (see Section 2.1 for PFAS-specific chemical properties, including predicted LogP), but accumulate in locations such as the blood, liver, and kidneys (and can be transferred to offspring through placental transfer and breast milk) (ATSDR, 2018; U.S. EPA, 2016c, d; Post et al., 2012). However, as illustrated in Table 7, previous summaries of the existing literature suggest there are pronounced half-life differences across species, sex, and type of PFAS. In general, PFAS with longer chain lengths are reported to have a longer serum half-life. For the PFAS with data available, serum half-life variation across species exhibits the following pattern: rats<mice<monkeys<humans. The extent of this cross-species difference appears to be greater than would be predicted by standard allometric (body-weight scaling to the ¾ power [$BW^{3/4}$]) scaling, and for mice versus rats is in the opposite direction (i.e., allometric scaling would predict mice<rats<monkeys<humans). Finally, except for PFDA (for which the least half-life information is available), the remaining PFAS being assessed are reported to have a shorter serum half-life in females than in males, sometimes markedly so (e.g., PFNA in rats). The approach to validating and possibly refining these values (e.g., based on new data) in these PFAS assessments is outlined in Section 9.2.1. Notably, it is not expected that there will be enough data to examine lifestage-specific differences in absorption, distribution, metabolism, and excretion (ADME).

The apparent toxicokinetic differences may have a significant impact on the interpretation of toxic effects across species or sexes. More directly, substantive toxicokinetic differences would be expected to affect quantitative extrapolations of dose-response data from experimental animals to humans. Thus, the half-life estimates for these five PFAS are likely to impact multiple assessment decisions, and a critical review of the available ADME data for each PFAS will be important (see discussion in Sections 5 and 9.2).

Although not identified during the preliminary literature inventory shown above, physiologically based pharmacokinetic (PBPK) models for PFHxS (Kim et al., 2018) and PFDA and PFNA (Kim et al., 2019) parameterized for adult male and female rats and humans have recently been described. Fàbrega et al. (2015) also described a PBPK model for multiple PFAS in humans, including PFBS, PFHxS, PFHxA, PFNA, and PFDA. In addition, Verner et al. (2016) adapted a classical pharmacokinetic model structure for evaluating gestational and lactational transfer of PFAS from mothers to their children, including PFHxS. These models could prove useful for addressing toxicokinetic questions in these assessments (see discussion in Sections 6.4 and 11.2).

Table 7. Preliminary serum half-life estimates of five per- and polyfluoroalkyl substances (PFAS) across species and sexes

	PFBA (C4)		PFHxA (C6)		PFHxS (C6)		PFNA (C9)		PFDA (C10)	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Rat	1.0–1.8 h	6–9 h	0.4–0.6 h	1.0–1.6 h	1.8 d	6.8 d	1.4 d	30.6 d	58.6 d	39.9 d
Mouse	3 h	12 h	~1.2 h	~1.6 h	24–27 d	28–30 d	26–68 d	34–69 d	ND	
Monkey	1.7 d		2.4 h	5.3 h	87 d	141 d	ND		ND	
Human	3 d		32 d		8.5 yr		4.3 yr		12 yr	

"C" = carbon chain length; ND = no data.

Data are summarized in Lau (2015). Note that these values do not necessarily represent those that would be used in qualitative or quantitative analyses for these PFAS assessments because the underlying data will be reviewed and possibly supplemented with additional (e.g., newer) studies. Darker shading indicates longer half-life (i.e., from hours to days to years).

2.4.2. Human Relevance of Effects in Animals that Involve Peroxisome Proliferator-Activated Receptor Alpha (PPARα) Receptors

Activation of the peroxisome proliferator-activated receptor alpha (PPARα) by PFAS has been reported, with *in vitro* evidence that the potency of human and mouse PPARα activation is positively correlated with increasing PFCA chain length up to C9 (no human receptor activation was noted for PFDA, although activation of the mouse receptor was only slightly less potent than PFNA) and greater for carboxylates than sulfonates (Wolf et al., 2014; Wolf et al., 2008; Takacs and Abbott, 2007; Shipley et al., 2004; Maloney and Waxman, 1999). It is not known whether PFAS distribute to the nucleus and bind directly to PPARα *in vivo*, or whether these substances activate the receptor indirectly. PPARα ligand binding causes a conformational change in the protein, release of corepressors, heterodimerization with the retinoid X receptor (RXR), and binding to cognate peroxisome proliferator response elements in the promoters of target genes (perhaps most notably, those related to fatty acid β-oxidation and energy homeostasis) to modulate gene transcription.

PPARα is a ligand-activated nuclear receptor expressed in many tissues and has been at the forefront of a longstanding debate as to whether chemical-induced PPARα modulation in rodents, particularly in the liver, is relevant to humans (Corton et al., 2018; Filgo et al., 2015; Guyton et al., 2009). PPARα is active in humans and responsive to the hypolipidemic effect of fibrate drugs that lower serum lipid levels, but the human receptor is generally considered less sensitive than PPARα in rodents (Corton et al., 2014; Wolf et al., 2014; Wolf et al., 2008; Maloney and Waxman, 1999). In particular, research with fibrate drugs indicates that PPARα activation in rodents leads to hepatic effects such as hepatomegaly, peroxisome proliferation, and in some instances over the long term, cancer. These effects are not observed in human models (Corton et al., 2014). Evaluating the human relevance of animal PPARα evidence is complicated by a lack of comparable model systems,

including widely used primary cell lines that rapidly lose the capability to express nuclear receptors such as PPARα (Soldatow et al., 2013) and potential species-specific differences in transcriptional coactivators and other pathway components.

With PFAS specifically, PPARα-dependent effects have been best studied in relation to hepatic effects (and liver cancer, in particular). However, PPARα is also known to be important to other physiological processes in both rodents and humans, including energy homeostasis, inflammation, reproduction, musculoskeletal function, and development (Corton et al., 2014; Burri et al., 2010; Abbott, 2009; Peraza et al., 2006; Corton et al., 2000). Thus, although not well studied for PFAS, the modulation of PPARα may be important to consider for developmental, metabolic, reproductive, and immunological effects, as well as for hepatic effects.

There are additional complexities to considering the dependence on, and human relevance of, PPARα activation by PFAS for certain health effects. The extent of PPARα activation is likely to differ by PFAS type, making it harder to apply read-across (specifically, drawing conclusions for one PFAS based on findings for another PFAS) or related approaches. In addition, based on conclusions from other PFAS assessments and review articles, there is evidence to indicate that many PFAS-mediated effects appear to include both PPARα-dependent and PPARα-independent mechanisms, the latter of which include activation of PPARγ, phosphatidylinositol-3-kinase-serine/threonine kinase Akt (PI3K-Akt), constitutive androstane receptor (CAR), nuclear factor kappa B pathway (NF-κB), farnesoid X receptor, liver X receptor, and estrogen receptor α (Li et al., 2017; Rosen et al., 2017; FSANZ, 2016; U.S. EPA, 2014b, d; Foreman et al., 2009).

Despite the complexities involved, it is important to evaluate the human relevance of some PFAS exposure-mediated effects in animals (see discussion in Section 9.2).

2.4.3. Potential Confounding by Other Per- and Polyfluoroalkyl Substances (PFAS) Exposures in Epidemiology Studies

Because different PFAS may be used in similar applications or result from similar sources, potential confounding of associations by PFAS coexposures is an important area of uncertainty for epidemiology studies. When associations are found for two or more moderately correlated PFAS in a study, including those not the focus of these assessments (e.g., PFOS and PFOA), confounding is a possible explanation. Based on a cursory review of studies identified during the preliminary literature inventory, a complicating factor is that correlations between PFAS pairs vary considerably across studies (see Section 6.2.1). When a study does not report the correlations in its population, the interpretation of the risk of bias from confounding is particularly challenging. Even when correlations are reported, there is no perfect method for eliminating confounding. Given this variability, assessing the likelihood and impact of this source of potential confounding based on reporting within individual studies is expected to be difficult (see discussion in Section 6.2).

2.4.4. Toxicological Relevance of Changes in Certain Urinary and Hepatic Endpoints in Rodents

The scientific community has identified difficulties in interpreting the toxicological relevance of changes in certain urinary and hepatic endpoints available in rodent studies (based on the preliminary literature inventory) for some of the five PFAS assessments. The specific rodent endpoints in question are chronic progressive nephropathy and related urinary histopathological changes (including alpha 2u-globulin-mediated changes), and hepatic effects that may be considered adaptive (e.g., increased liver weight; cellular hypertrophy; single cell necrosis/apoptosis). For the former, some of these changes are not considered relevant to humans, and methods exist for evaluating the dependency of observed changes on this rodent-specific mechanism. For the latter, neither a clear scientific consensus nor specific EPA-wide guidance defines exactly what level of change or constellation of effects is necessary to establish cause for concern. Thus, interpretations of the toxicological relevance of changes in these specific endpoints are expected to require additional consideration (see discussion in Section 9.2).

2.4.5. Characterizing Uncertainty Due to Missing Chemical-Specific Information

Two PFAS, PFOA and PFOS (C8), have been studied more extensively than other PFAS. Thus, this existing knowledge base may be useful in helping to characterize existing data gaps and uncertainties in the current five PFAS assessments. Two recently developed EPA assessments (PFBS and GenX chemicals) could also provide information during the development of these current assessments. For example, given knowledge regarding the health effects of PFOA and PFOS, the potential lack of studies on immune effects for PFBA and developmental effects for PFHxA (based on the preliminary literature inventory; see Section 2.3.2) appear to represent important database uncertainties. In addition, given the potential for lifetime human exposure to PFAS by multiple routes of exposure (see Section 2.1.5), the apparent scarcity of data on most of these five PFAS other than short-term oral exposure studies in animals is expected to affect assessment decisions and characterization of uncertainties (see discussion in Sections 10.2 and 11.2.3).

3. OVERALL OBJECTIVES, SPECIFIC AIMS, AND POPULATIONS, EXPOSURES, COMPARATORS, AND OUTCOMES (PECO) CRITERIA

The overall objective of these five assessments is to identify adverse human health effects and characterize exposure-response relationships for the effects of perfluorobutanoic acid (PFBA), perfluorohexanoic acid (PFHxA), perfluorohexanesulfonate (PFHxS), perfluorononanoic acid (PFNA), and perfluorodecanoic acid (PFDA) to support development of toxicity values. These assessments will use systematic review methods to evaluate the epidemiological and toxicological literature, including consideration of relevant mechanistic evidence (e.g., to inform key science issues; see Section 2.4). The evaluations conducted in these assessments will be consistent with relevant Environmental Protection Agency (EPA) guidance.[2]

The specific approach taken for these assessments of the potential health effects of PFBA, PFHxA, PFHxS, PFNA, and PFDA (and their associated salts) was based on input received during scoping, as well as a preliminary literature inventory of the health effects studied for these PFAS. As outlined in Section 2.3.2, these assessments will evaluate the potential for PFAS exposure via the oral or inhalation route to cause health effects in humans, specifically focusing on developmental effects; endocrine (primarily thyroid hormone) effects; hepatic effects, including lipid and lipoprotein measures; immune effects; reproductive effects in males or females; urinary effects; general toxicity; and carcinogenicity (see Section 5 for preliminary decisions for grouping outcomes and endpoints within each of these predetermined health effect categories). Data on cardiovascular effects, hematological effects, metabolic effects including diabetes, and nervous system effects will also be summarized when available. These summaries may be developed in association with one of the health effects noted above either as a separate formal evaluation of hazard or as part of a qualitative summary on "other effects," depending on the assessment-specific data. Given the paucity of available studies and in the absence of exceptional evidence in any available studies, information on other health effects (i.e., gastrointestinal effects; musculoskeletal effects; ocular effects; and respiratory effects) will not be formally evaluated (these effects may be briefly summarized) in any of these assessments; although new literature relating to these outcomes will be monitored during literature search updates for potential inclusion. As outlined in the EPA PFAS action plan,[3] the characterization of the potential human health hazards from exposure to these

[2]EPA guidance documents: http://www.epa.gov/iris/basic-information-about-integrated-risk-information-system#guidance/.
[3]EPA PFAS action plan: https://www.epa.gov/pfas/epas-pfas-action-plan.

individual PFAS will be coupled with data generated from new advances in computational and high-throughput toxicology to inform evaluations of other PFAS.

3.1. SPECIFIC AIMS

The aims of these assessments are to:

- Identify epidemiological (i.e., human) and toxicological (i.e., experimental animal) literature reporting effects of exposure to PFBA, PFHxA, PFHxS, PFNA, and PFDA (and their associated salts), as outlined in the PECO. These five systematic reviews will focus on identifying studies following oral or inhalation exposure to PFAS.

- Evaluate mechanistic information (including toxicokinetic understanding) associated with exposure to PFBA, PFHxA, PFHxS, PFNA, and PFDA, to inform the interpretation of findings related to potential health effects in studies of humans and animals. The scope of these analyses of mechanistic information will be determined by the complexity and confidence in the phenotypic evidence in humans and animals, the likelihood of the analyses (e.g., considering the mechanistic studies available based on the literature inventory; see Section 4.2.2) to affect evidence synthesis conclusions for human health, and the directness or relevance of the available model systems for understanding potential human health hazards (see Section 9.2). The mechanistic evaluations will focus primarily on the key science issues identified in Section 2.4.

- Conduct study evaluations for individual epidemiological and toxicological studies (evaluating reporting quality, risk of bias, and sensitivity) and PBPK) (scientific and technical review). The evaluation of epidemiology studies will specifically consider, to the extent possible, the likelihood and impact of potential confounding by other PFAS (see Section 6.2.1).

- Extract data on relevant health outcomes from epidemiological and toxicological studies of high, medium, and low confidence based on the study evaluations (full data extraction of low confidence studies may not be performed for poorly studied health effects or for health effects on which extensive *medium* and *high* confidence studies exist in the evidence base).

- Synthesize the evidence across studies, assessing similar health outcomes using a narrative approach. To inform future comparisons across a range of PFAS structures and properties (e.g., using high throughput screening, computational toxicology approaches, and chemical informatics to fill in data gaps; see EPA PFAS action plan), each of the five PFAS assessments will synthesize the available evidence (or lack thereof) for developmental effects; endocrine (primarily thyroid hormone) effects; hepatic effects, including lipid and lipoprotein measures (the latter of which are also applicable to interpreting the potential for cardiovascular toxicity); immune effects; reproductive effects in males or females; urinary effects; general toxicity; and carcinogenicity. Some assessments may include additional evidence syntheses for other health effects. The toxicological relevance of changes in some urinary and hepatic outcomes will be a point of focus in the evidence syntheses (see Section 9.2.3).

- For each health outcome (or grouping of outcomes), evaluate the strength-of-evidence across studies (or subsets of studies) separately for studies of exposed humans and for animal studies. Based on the focused mechanistic analyses specific to each PFAS assessment (see Section 9.2), the mechanistic evidence will be used to inform evaluations of the available health effects evidence (or lack thereof).

- For each health outcome (or grouping of outcomes), develop an integrated expert judgment across evidence streams as to whether the evidence is sufficient (or insufficient) to indicate that exposure to the PFAS has the potential to be hazardous to humans (in rare instances, the evidence may be judged as sufficient to indicate that a hazard is unlikely). The judgment will be directly informed by the evidence syntheses and based on structured review of an adapted set of considerations for causality first introduced by Austin Bradford Hill (Hill, 1965)(see Sections 9 and 10), including consideration (e.g., based on available mechanistic information) and discussion of biological understanding. As part of the evidence integration narrative, characterize the strength of evidence for the available database of studies and its uncertainties, and identify and discuss issues concerning potentially susceptible populations and lifestages.

- Derive toxicity values (e.g., oral reference doses [RfDs], inhalation reference concentrations [RfCs], cancer risk estimates) as supported by the available data (see Section 10.2). Apply toxicokinetic and dosimetry modeling (possibly including PBPK modeling) to account for interspecies differences, as appropriate. Given the apparent species and sex differences in the toxicokinetic profile of the different PFAS (see Section 2.4), methods to address these potential differences will be a key consideration (see Section 9.2.1).

- Characterize uncertainties and identify key data gaps and research needs across each PFAS database, such as limitations of the available evidence, limitations of the systematic review, and consideration of dose relevance and toxicokinetic differences when extrapolating findings from higher dose animal studies to lower levels of human exposure.

3.2. POPULATIONS, EXPOSURES, COMPARATORS, AND OUTCOMES (PECO) CRITERIA

The PECO criteria are used to identify the evidence that addresses the specific aims of the assessment and to focus the literature screening, including study inclusion/exclusion, in a systematic review (see details on literature screening in Section 4.2). Given the anticipated lack of studies on carcinogenicity for these PFAS based on the preliminary literature inventory, genotoxicity studies were included in the PECO criteria (see Table 8).

In addition to those studies meeting the PECO criteria, studies containing supplemental material that are potentially relevant to the specific aims of the assessment were tracked during the literature screening process. Although these studies did not meet PECO criteria, they were not excluded from further consideration. The categories used to track studies as "potentially relevant supplemental material" are also described in Section 4.2.

Table 8. Populations, exposures, comparators, and outcomes (PECO) criteria

PECO element	Evidence
Populations	Human: Any population and lifestage (occupational or general population, including children and other sensitive populations). The following study designs will be included: controlled exposure, cohort, case-control, and cross-sectional. (Note: Case reports and case series will be tracked as potential supplemental material.) Animal: Nonhuman mammalian animal species (whole organism) of any lifestage (including preconception, *in utero*, lactation, peripubertal, and adult stages). Other: *In vitro*, *in silico*, or nonmammalian models of genotoxicity. (Note: Other *in vitro*, *in silico*, or nonmammalian models will be tracked as potential supplemental material.)
Exposures	Human: Studies providing quantitative estimates of PFAS exposure based on administered dose or concentration, biomonitoring data (e.g., urine, blood, or other specimens), environmental or occupational-setting measures (e.g., water levels or air concentrations, residential location and/or duration, job title, or work title). (Note: Studies that provide qualitative, but not quantitative, estimates of exposure will be tracked as supplemental material.) Animal: Oral or Inhalation studies including quantified exposure to a PFAS of interest based on administered dose, dietary level, or concentration. (Note: Non-oral and non-inhalation studies will be tracked as potential supplemental material.) PFAS mixture studies are included if they employ an experimental arm that involves exposure to a single PFAS of interest. (Note: Other PFAS mixture studies are tracked as potential supplemental material.) Studies must address exposure to one or more of the following: PFDA (CASRN 335-76-2), PFDA ammonia salt (CASRN 3108-42-7), PFDA sodium salt (CASRN 3830-45-3), PFNA (CASRN 375-95-1), PFNA ammonium salt (CASRN 4149-60-4), PFNA sodium salt (CASRN 21049-39-8), PFHxA (CASRN 307-24-4), PFHxA sodium salt (CASRN 2923-26-4), PFHxA ammonium salt (CASRN 21615-47-4), PFHxS (CASRN 355-46-4), PFHxS potassium salt (CASRN 3871-99-6), PFBA (CASRN 375-22-4), or PFBA ammonium salt (CASRN 10495-86-0). [Note: although most PFAS tend to remain unchanged in the body (Nabb et al., 2007), it is possible that some PFAS may be bio-transformed to a PFAS of interest. Thus, studies of precursor PFAS that identify and quantify a PFAS of interest (e.g., as a metabolite) will be tracked as potential supplemental material (e.g., for ADME interpretations).]
Comparators	Human: A comparison or reference population exposed to lower levels (or no exposure/exposure below detection levels) or for shorter periods of time. Animal: Includes comparisons to historical controls or a concurrent control group that is unexposed, exposed to vehicle-only or air-only exposures. (Note: Experiments including exposure to PFAS across different durations or exposure levels without including one of these control groups will be tracked as potential supplemental material [e.g., for evaluating key science issues; Section 2.4].)
Outcomes	All cancer and noncancer health outcomes. (Note: Other than genotoxicity studies, studies including only molecular endpoints [e.g., gene or protein changes; receptor binding or activation] or other non-phenotypic endpoints addressing the potential biological or chemical progression of events contributing towards toxic effects will be tracked as potential supplemental material [e.g., for evaluating key science issues; Section 2.4].)

PECO element	Evidence
PBPK models	Studies describing physiologically based pharmacokinetic (PBPK) and other PK models for PFDA (CASRN 335-76-2), PFDA ammonia salt (CASRN 3108-42-7), PFDA sodium salt (CASRN 3830-45-3), PFNA (CASRN 375-95-1), PFNA ammonium salt (CASRN 4149-60-4), PFNA sodium salt (CASRN 21049-39-8), PFHxA (CASRN 307-24-4), PFHxS (CASRN 355-46-4), PFHxS potassium salt (CASRN 3871-99-6), PFBA (CASRN 375-22-4), or PFBA ammonium salt (CASRN 10495-86-0).

ADME = absorption, distribution, metabolism, and excretion; PK = pharmacokinetic.

4. LITERATURE SEARCH AND SCREENING STRATEGIES

The initial literature search was completed in July 2017 as part of a cross-Environmental Protection Agency (EPA) workgroup that focused on a large set PFAS, including the five PFAS addressed in this protocol. Subsequent literature searches were refined and are being updated regularly. In an effort to ensure that all pertinent studies were captured, the studies identified as relevant to PFBA, PFHxA, PFHxS, PFNA, and PFDA are being shared simultaneously with release of this protocol.[4] These search efforts reflect studies published through February 2018, several unpublished reports, and a few recent studies not yet identified through the formal literature updating process; the literature is currently being updated and will be updated regularly until several months before public release of the draft assessments.[5] Accordingly, the methods for literature search and screening (as well as some of the approaches to refining the evaluation plan based on the identified literature; see Section 5) are described in the protocol using the past tense, whereas approaches for the other assessment methods are outlined using the future tense.

4.1. LITERATURE SEARCH STRATEGIES

The initial literature search strategy performed in July 2017 was designed to identify a broad range of topics relevant to PFAS, including studies on physicochemical properties, environmental fate and occurrences, human exposures, and biological effects representative of all types of evidence (i.e., human, animal, *in vitro*, *in silico*) and health outcomes. PFAS search terms included PFAS names (including salt, cationic, and anionic forms), all known synonyms, and CAS registry numbers. The literature search itself encompassed a non-date-limited query of the following databases:

- PubMed (National Library of Medicine)
- Web of Science (Thomson Reuters)

[4]PFBA: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2632
PFHxA: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2628
PFHxS: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2630
PFNA: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2633
PFDA: https://hero.epa.gov/hero/index.cfm/project/page/project_id/2614

[5]Although not identified (yet) as part of the formal literature searches, several recent PBPK studies found through regular monitoring of new studies are included in this protocol (see Section 6.4) so that the process for evaluating those data can be outlined.

- Toxline (National Library of Medicine)
- TSCATS (Toxic Substances Control Act Test Submissions)

All literature identified in the initial search was loaded into the U.S. EPA Health and Environmental Research Online (HERO) database. In February 2018, the literature search was updated for the PFAS in this assessment (i.e., PFBA, PFHxA, PFHxS, PFNA, and PFDA). The updated literature query included all PFAS nomenclature from the initial search as well as a broader non-date-limited search of several new PFAS synonyms that had been identified since the original search. This updated search was conducted by EPA's HERO tool to search the same databases as were included in the initial literature query.

Because each database has its own search architecture, the resulting search strategy was tailored to account for each database's unique search functionality. Full details of the July 2017 and February 2018 search strategies are presented in Appendix B. No literature was restricted by language.

Additional relevant literature not found through database searching was identified by:

- Review of studies cited in state, national (EPA, Food and Drug Administration [FDA], etc.), and international (International Agency for Research on Cancer [IARC], World Health Organization [WHO], European Chemicals Agency [ECHA], etc.) assessments on these five PFAS, including parallel assessment efforts in progress (e.g., the draft Agency for Toxic Substances and Disease Registry [ATSDR] assessment released publicly in 2018).
- Review of studies submitted to federal regulatory agencies and brought to the attention of EPA. For example, studies submitted to EPA by the manufacturers of these five PFAS in support of requirements under the Toxic Substances Control Act (TSCA). (Note: such studies [or data summaries] will only be tracked in the literature flow diagrams released with each of the five assessments when they are going to be made publicly available.)
- Identification of studies during screening for other PFAS. For example, epidemiology studies relevant to more than one of these five PFAS were sometimes identified by searches focused on one PFAS, but not the others.
- Other gray literature (i.e., primary studies not indexed in typical databases, such as technical reports from government agencies or scientific research groups; unpublished laboratory studies conducted by industry; or working reports/white papers from research groups or committees) brought to the attention of EPA during problem formulation, engagement with technical PFAS experts, and during future solicitation of Agency, interagency, and public comment during the Integrated Risk Information System (IRIS) assessment development and review process. For example, one such study was brought to the attention of the EPA on March 29, 2018 by the National Toxicology Program (NTP) when NTP published study tables and individual animal data from a 28-day toxicity study of multiple PFAS (NTP, 2011). A peer-reviewed NTP Technical Report was not yet available at the time this protocol was drafted, but these data have undergone standard NTP quality

assurance/control processing and are publicly available, and a protocol outlining the NTP study methods is available in HERO (https://hero.epa.gov/hero/index.cfm/reference/details/reference_id/4309741).

The number of studies on PFBA, PFHxA, PFHxS, PFNA, and PFDA returned from the literature searches through February 2018 is documented in the literature flow diagrams in Figure 4, which also reflect the literature screening decisions (see Section 4.2). Notably, the identification and review of records submitted to the EPA, which may include confidential business information (CBI), is ongoing. This includes exploring the possibility of making the data within any identified records publicly available. Any records included in the assessments will be reflected in updates to the protocol and in the draft assessments. In addition, any identified companion documents for the included studies, such as retractions, corrections and supplemental materials, will also be included, and the assessments will incorporate the most recent publication materials (note: these are tracked as separate, "included" records in the literature flow diagrams [see Section 4.2.2] and HERO; companion documents in other screening categories such as "excluded," which are not relevant to the target PECO, are similarly tagged as separate records within that screening category).

The literature searches will be updated throughout the assessments' development and review process to identify newly published literature. The last full literature search update will be conducted prior to (several months) the planned release of the draft document for public comment. The literature flow diagrams (see Section 4.2.2) will be updated with the results of these updates. Although uncommon, it is possible that during assessment development and review additional literature searches may be performed (e.g., to supplement an analysis of a specific mechanism or biological linkage). Any such ancillary searches will be documented in updates to the protocol.

The IRIS Program takes extra steps to ensure identification of pertinent studies by encouraging the scientific community and the public to identify additional studies and ongoing research; by searching for publicly available data submitted under the TSCA and the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA); and by considering late-breaking studies that would affect the credibility of the conclusions, even during the review process. Studies identified after peer review begins will only be considered for inclusion if they are directly relevant to the PECO criteria and are expected to fundamentally alter the assessment's conclusions. Release of the PECO-screened literature in parallel with release of the protocol for public comment provides an opportunity for stakeholders to identify any missing studies, which if identified, will be screened as outlined above for adherence to the PECO criteria.

4.1.1. Non-Peer-Reviewed Data

IRIS assessments rely mainly on publicly accessible, peer-reviewed studies. However, it is possible that gray literature (i.e., studies that are not reported in the peer-reviewed literature) directly relevant to the PECO may be identified during assessment development (e.g., good

laboratory practices [GLP] studies submitted to EPA, dissertations, etc.). In this case, if the data make a substantial impact on assessment decisions or conclusions (i.e., have potential to affect the PECO statement, hazard conclusions, or dose-response analysis), EPA can obtain external peer review if the owners of the data are willing to have the study details and results made publicly accessible. This independent, contractor-driven peer review would include an evaluation of the study, as is done for peer review of a journal publication. The contractor would identify and select two to three scientists knowledgeable in scientific disciplines relevant to the topic as potential peer reviewers. Persons invited to serve as peer reviewers would be screened for conflict of interest before confirming their service. In most instances, the peer review would be conducted by letter review. The study authors would be informed of the outcome of the peer review and given an opportunity to clarify issues or provide missing details. The study and its related information, if used in the IRIS assessment, would become publicly available. In the assessment, EPA would acknowledge that the document underwent external peer review managed by the EPA, and the names of the peer reviewers would be identified. In certain cases, IRIS will conduct an assessment for utility and data analysis based on having access to a description of study methods and raw data that have undergone rigorous quality assurance/quality control review (e.g., ToxCast/Tox21 data; results of National Toxicology Program [NTP] studies) but that have not yet undergone external peer-review.

Unpublished (e.g., raw) data from personal author communication can supplement a peer-reviewed study, as long as that information is made publicly available. If such ancillary information is acquired, it will be documented in either the Health Assessment Workspace Collaborative (HAWC) or HERO project page for the PFAS being assessed (depending on the nature of the information received).

4.2. SCREENING PROCESS

As described below, PECO criteria or predefined inclusion and exclusion criteria (i.e., the latter were used for the initial search) were used by two independent reviewers to screen and inventory studies at the title and abstract level. For those studies considered relevant at the title and abstract level, these criteria were then used to determine inclusion or exclusion of a reference based on the full text. In addition to the PECO criteria, the following exclusion criteria were applied:

- Review, commentary, other agency assessment, letter, or other record that does not contain original data (note that these records were tracked for potential use in identifying study-specific, original data relevant to specific scientific questions during assessment development, including scanning of reference lists for unidentified studies; any such studies incorporated into the assessment will be tracked under "other" as the reference source in updates to the protocol)

- Study available only as an abstract (e.g., conference abstract)
- Full text of the study is not available, and screening decisions could not be made at the title/abstract level

In addition to including studies that meet PECO criteria, other studies containing material that is potentially relevant to the assessments' objectives and specific aims were tracked during the screening process as "potentially relevant supplemental material." These studies were not excluded, but they may not be incorporated into the assessments unless they are deemed to be relevant to addressing the key science issues, specific aims (see Sections 2.4 and 3.1), or key scientific uncertainties identified at later stages of assessment development (see Section 9). Studies categorized as "potentially relevant supplemental material" include the following:

- *In vivo* mechanistic or mode-of-action studies, including non-PECO routes of exposure and populations (e.g., nonmammalian models) and studies examining potential susceptibility
- *In vitro* and *in silico* models
- ADME and toxicokinetic studies (excluding models)
- Exposure assessment or characterization (no health outcome) studies
- PFAS mixture studies (no individual PFAS comparisons)
- Human case reports or case-series studies
- Ecotoxicity studies
- Studies on PFAS manufacture/use
- Treatment/remediation studies
- Studies of PFAS analysis or other laboratory methods
- Environmental fate and transport studies
- Studies of other PFAS

Several of these categories of studies were further screened for consideration in addressing the key science issues (described in Section 98).

Title and abstract screening. Following a pilot phase to calibrate screening guidance, two screeners independently performed a title and abstract screen using a structured form in DistillerSR (Evidence Partners; https://distillercer.com/products/distillersr-systematic-review-software/). For citations with no abstract, the article was excluded if screening decisions could not

be made based on the title and other citation information (e.g., page length) and additional attempts to acquire the abstract or full text were unsuccessful. Screening conflicts were resolved by discussion among the primary screeners with consultation by a third reviewer or technical advisor (if needed) to resolve any remaining disagreements. Eligibility status of non-English studies was assessed using the same approach with online translation tools used as needed to evaluate portions of the study text and assess eligibility at the title and abstract level.

Studies not meeting title/abstract criteria but identified as "potentially relevant supplemental material" were categorized (i.e., tagged) during the title and abstract screening process (further described in Section 4.3). Conflict resolution was not required during the screening process to identify supplemental information (i.e., tagging by a single screener was considered adequate to identify the study as potentially relevant supplemental material for possible inclusion during draft development).

Before beginning the Integrated Risk Information System (IRIS) PFAS assessments project, the EPA contractor that conducted the July 2017 literature search as part of an EPA-wide workgroup had performed a title and abstract screen to bin studies into different categories (e.g., human, *in vivo* animal, excluded). At the initiation of these PFAS assessments within IRIS, a formalized effort was deployed with a new title and abstract screen of all studies identified in the initial July 2017 search based on the PECO criteria in Table 8. For this initial literature screening, specific inclusion/exclusion criteria were applied in the formalized title and abstract screen (see Appendix B, Table B-6). Title and abstract screening of studies identified during literature search updates will be conducted using the PECO criteria in Table 8 in DistillerSR using forms that facilitate simultaneous initial tagging during screening (e.g., category of supplemental data; contains data on other PFAS of interest). An example of the questions and answers populating the DistillerSR form for title/abstract and full-text (below) screening during literature search updates is provided in Appendix B, Table B-7.

Full-text screening. Records that were not excluded based on the title and abstract were advanced to full-text review. Full-text copies of these potentially relevant records were retrieved, stored in the HERO database, and independently assessed by two screeners using a structured form in DistillerSR to confirm eligibility. Screening conflicts were resolved by discussion among the primary screeners with consultation by a third reviewer or technical advisor (as needed to resolve any remaining disagreements). As with the title and abstract screening, some studies were also identified as "potentially relevant supplemental material" based on full-text screening. Approaches for language translation included engagement of a native speaker from within EPA or use of fee-based translation services.

In addition to identifying studies as included, excluded, or potential supplemental material, the reviewers used the DistillerSR screening forms to confirm the specific PFAS (or multiple PFAS) evaluated and to document several important experimental features of the studies (see Section 4.3).

The results of this screening process are documented in the HERO database (https://hero.epa.gov; see Section 4.1 for links to the specific HERO pages) and literature flow diagrams (see Figure 4), with individual studies "tagged" in HERO according to their appropriate category descriptors (e.g., reference source; screening decisions regarding inclusion, exclusion, or identification as supplemental; type of study).

4.2.1. Multiple Publications of the Same Data

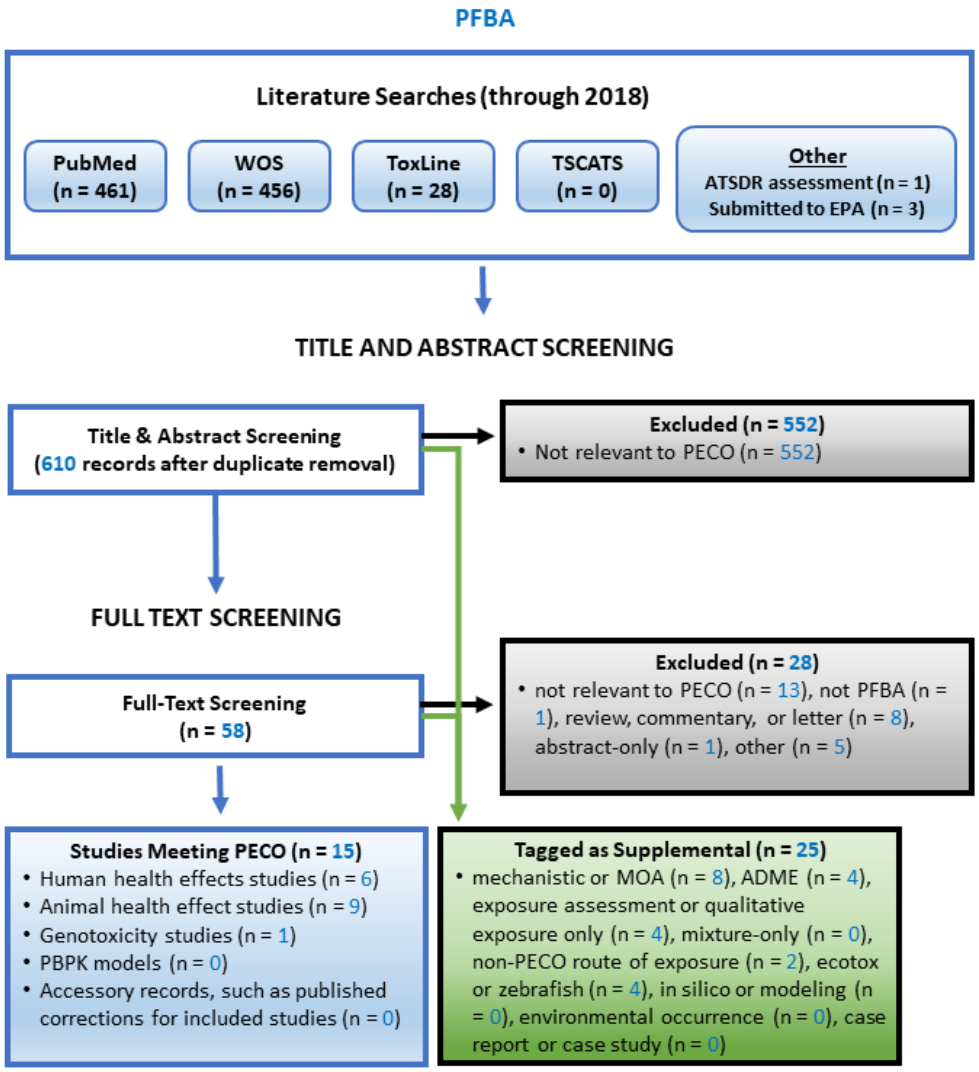
When there are multiple publications using the same or overlapping data, all publications on the research were included, with one selected for use as the primary study; the others were considered as secondary publications with annotation in HAWC indicating their relationship to the primary record during data extraction. For epidemiology studies, the primary publication was generally the one with the longest follow-up, the largest number of cases, or the most recent publication date. For animal studies, the primary publication was typically the one with the most recent publication date, longest duration of exposure, or the one that assessed the outcome(s) most informative to the PECO. For both epidemiology and animal studies, the assessments will include relevant data from all publications of the study, although if the same data are reported in more than one study, the data will only be extracted once (see Section 8). For corrections, retractions, and other companion documents to the included publications, a similar approach to annotation was taken (see Section 4.1), and the most recently published data will be incorporated in the assessments.

4.2.2. Literature Flow Diagrams

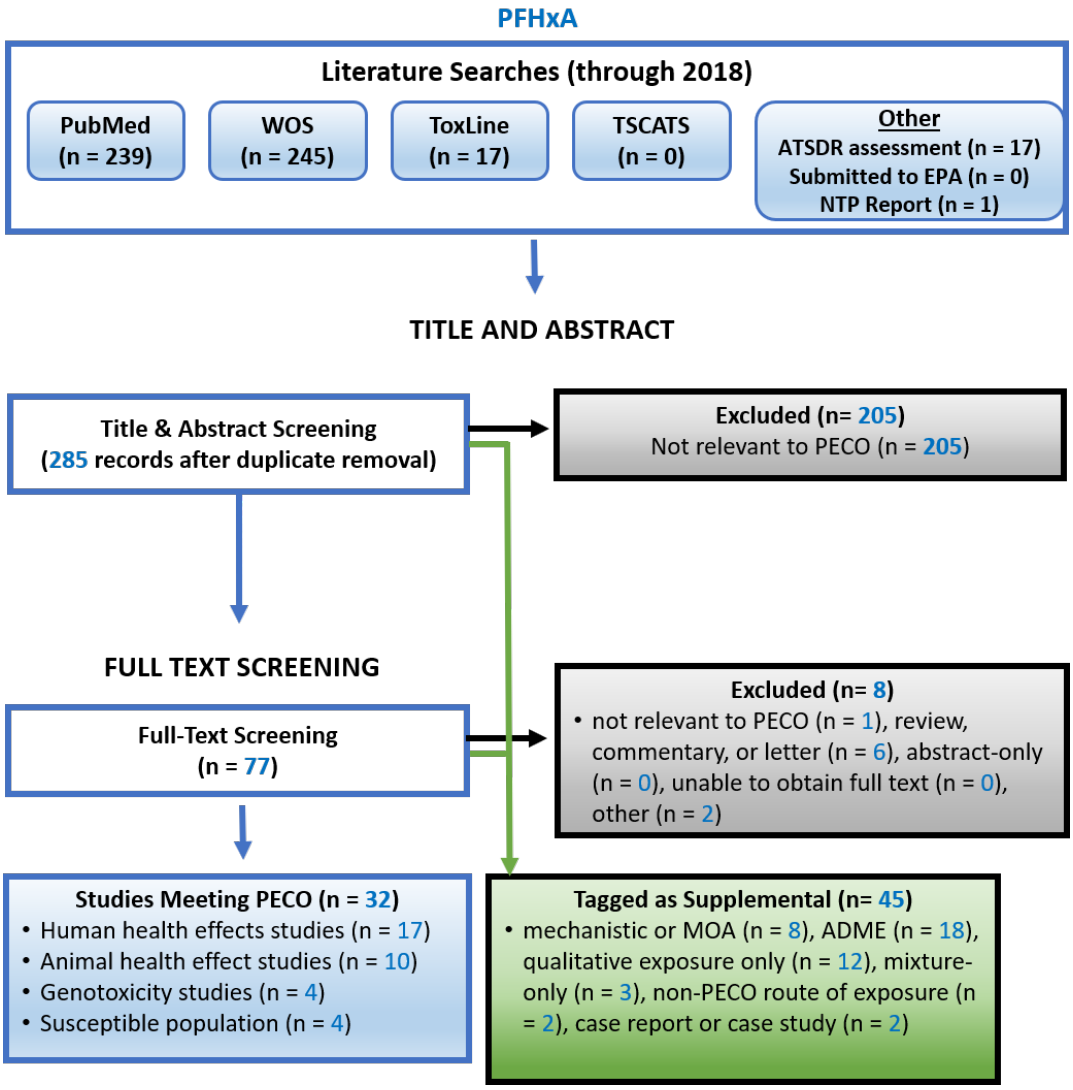
Figure 4 presents the literature flow diagrams for PFBA (a), PFHxA (b), PFHxS (c), PFNA (d), and PFDA (e).[6] These figures reflect literature searches through 2018. A literature search update has been conducted and the results will be reflected in the draft assessments (and the most current results can be viewed at any time in the HERO project pages provided in Section 4.1). Note that the potential for updates or revisions to these figures related to CBI data and other reference decisions is discussed in the previous sections.

[6]Note that the literature searches included the associated salts for each of the five PFAS, as presented in Figure 1 (Section 2.1.1). In addition, although not identified (yet) as part of the formal literature searches and not included in these diagrams, several recent PBPK studies found through regular monitoring of new studies are included in this protocol (see Section 6.4) so that the process for evaluating those studies can be outlined.

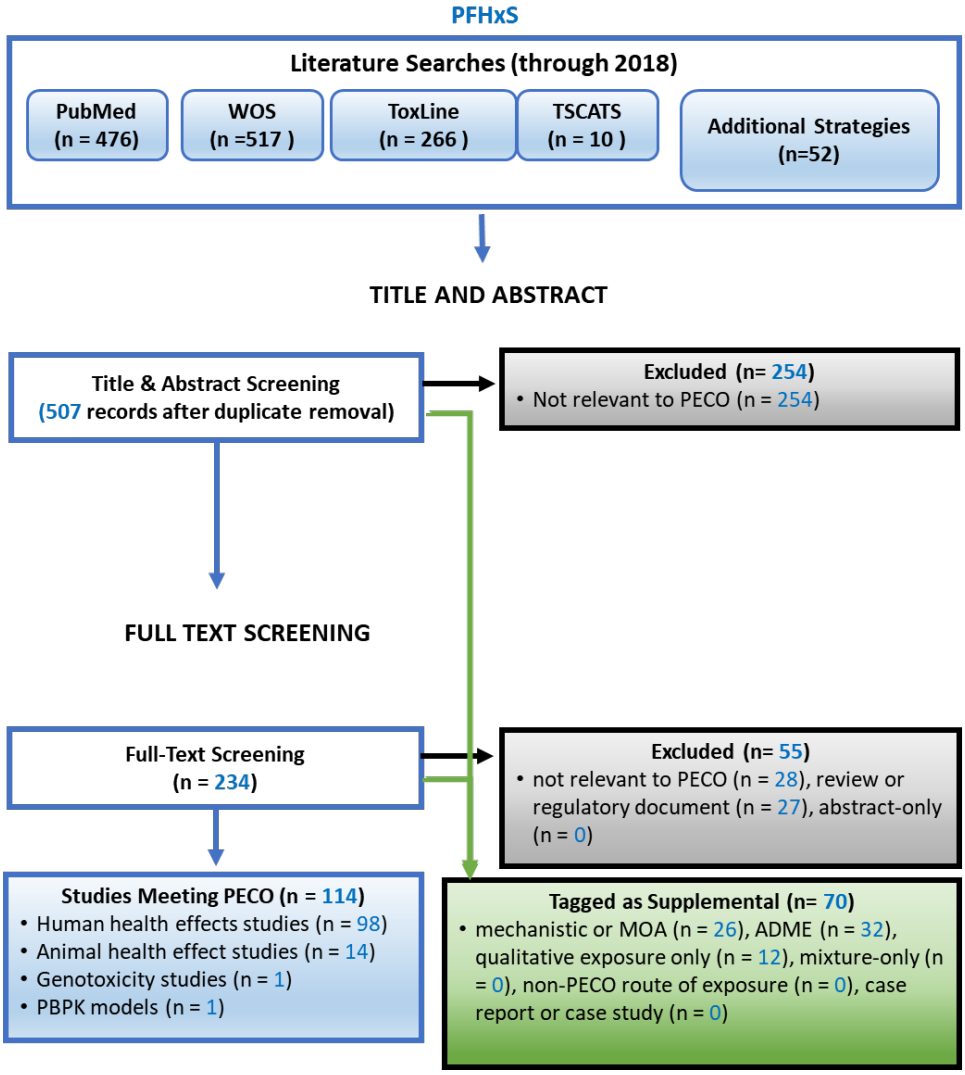
(a)


PFBA
Literature Searches (through 2018)
PubMed (n = 461)
WOS (n = 456)
ToxLine (n = 28)
TSCATS (n = 0)
Other
ATSDR assessment (n = 1)
Submitted to EPA (n = 3)
TITLE AND ABSTRACT SCREENING
Title & Abstract Screening (610 records after duplicate removal)
Excluded (n = 552)
• Not relevant to PECO (n = 552)
FULL TEXT SCREENING
Full-Text Screening (n = 58)
Excluded (n = 28)
• not relevant to PECO (n = 13), not PFBA (n = 1), review, commentary, or letter (n = 8), abstract-only (n = 1), other (n = 5)
Studies Meeting PECO (n = 15)
• Human health effects studies (n = 6)
• Animal health effect studies (n = 9)
• Genotoxicity studies (n = 1)
• PBPK models (n = 0)
• Accessory records, such as published corrections for included studies (n = 0)
Tagged as Supplemental (n = 25)
• mechanistic or MOA (n = 8), ADME (n = 4), exposure assessment or qualitative exposure only (n = 4), mixture-only (n = 0), non-PECO route of exposure (n = 2), ecotox or zebrafish (n = 4), in silico or modeling (n = 0), environmental occurrence (n = 0), case report or case study (n = 0)

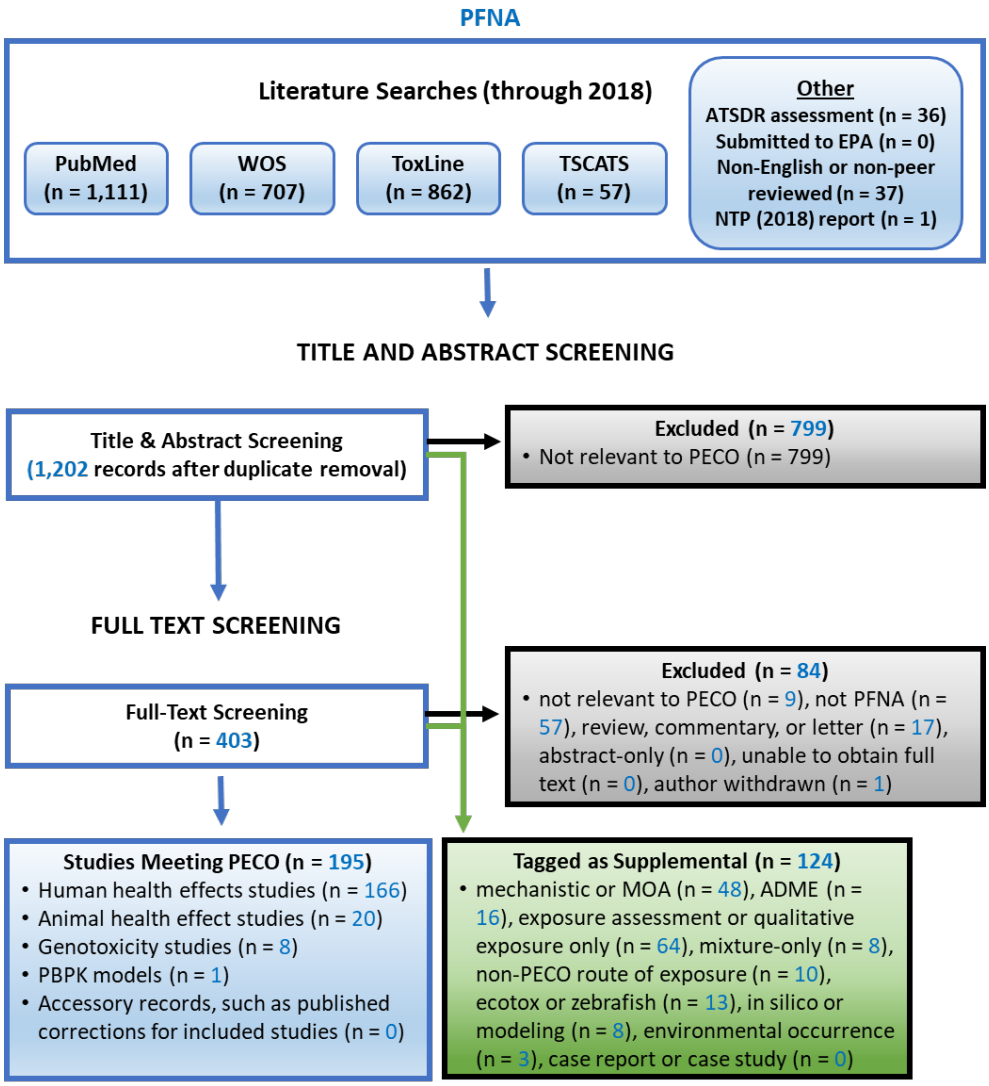
(b)


PFHxA
Literature Searches (through 2018)
PubMed (n = 239)
WOS (n = 245)
ToxLine (n = 17)
TSCATS (n = 0)
Other
ATSDR assessment (n = 17)
Submitted to EPA (n = 0)
NTP Report (n = 1)
TITLE AND ABSTRACT
Title & Abstract Screening (285 records after duplicate removal)
Excluded (n= 205)
Not relevant to PECO (n = 205)
FULL TEXT SCREENING
Full-Text Screening (n = 77)
Excluded (n= 8)
• not relevant to PECO (n = 1), review, commentary, or letter (n = 6), abstract-only (n = 0), unable to obtain full text (n = 0), other (n = 2)
Studies Meeting PECO (n = 32)
• Human health effects studies (n = 17)
• Animal health effect studies (n = 10)
• Genotoxicity studies (n = 4)
• Susceptible population (n = 4)
Tagged as Supplemental (n= 45)
• mechanistic or MOA (n = 8), ADME (n = 18), qualitative exposure only (n = 12), mixture-only (n = 3), non-PECO route of exposure (n = 2), case report or case study (n = 2)

(c)


PFHxS
Literature Searches (through 2018)
PubMed (n = 476)
WOS (n =517)
ToxLine (n = 266)
TSCATS (n = 10)
Additional Strategies (n=52)
TITLE AND ABSTRACT
Title & Abstract Screening (507 records after duplicate removal)
Excluded (n= 254)
• Not relevant to PECO (n = 254)
FULL TEXT SCREENING
Full-Text Screening (n = 234)
Excluded (n= 55)
• not relevant to PECO (n = 28), review or regulatory document (n = 27), abstract-only (n = 0)
Studies Meeting PECO (n = 114)
• Human health effects studies (n = 98)
• Animal health effect studies (n = 14)
• Genotoxicity studies (n = 1)
• PBPK models (n = 1)
Tagged as Supplemental (n= 70)
• mechanistic or MOA (n = 26), ADME (n = 32), qualitative exposure only (n = 12), mixture-only (n = 0), non-PECO route of exposure (n = 0), case report or case study (n = 0)

(d)


PFNA
Literature Searches (through 2018)
PubMed (n = 1,111)
WOS (n = 707)
ToxLine (n = 862)
TSCATS (n = 57)
Other
ATSDR assessment (n = 36)
Submitted to EPA (n = 0)
Non-English or non-peer reviewed (n = 37)
NTP (2018) report (n = 1)
TITLE AND ABSTRACT SCREENING
Title & Abstract Screening (1,202 records after duplicate removal)
Excluded (n = 799)
• Not relevant to PECO (n = 799)
FULL TEXT SCREENING
Full-Text Screening (n = 403)
Excluded (n = 84)
• not relevant to PECO (n = 9), not PFNA (n = 57), review, commentary, or letter (n = 17), abstract-only (n = 0), unable to obtain full text (n = 0), author withdrawn (n = 1)
Studies Meeting PECO (n = 195)
• Human health effects studies (n = 166)
• Animal health effect studies (n = 20)
• Genotoxicity studies (n = 8)
• PBPK models (n = 1)
• Accessory records, such as published corrections for included studies (n = 0)
Tagged as Supplemental (n = 124)
• mechanistic or MOA (n = 48), ADME (n = 16), exposure assessment or qualitative exposure only (n = 64), mixture-only (n = 8), non-PECO route of exposure (n = 10), ecotox or zebrafish (n = 13), in silico or modeling (n = 8), environmental occurrence (n = 3), case report or case study (n = 0)

(e)

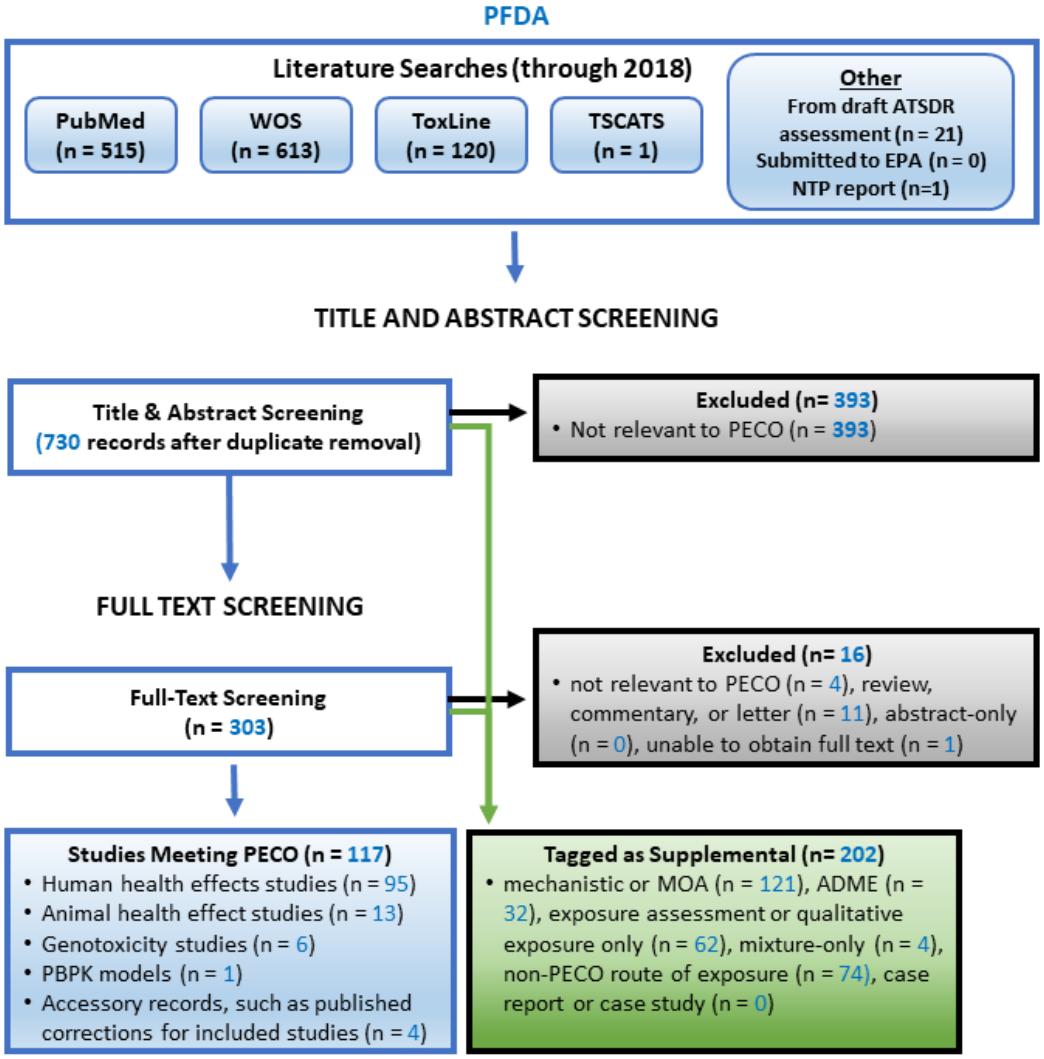


Figure 4. Literature flow diagrams for PFBA and its ammonium salt (a), PFHxA and its ammonium and sodium salts (b), PFHxS and its potassium salt (c), PFNA and its ammonium and sodium salts (d), and PFDA and its ammonium and sodium salts (e).

4.3. SUMMARY-LEVEL LITERATURE INVENTORIES

As noted in Section 4.2, during title/abstract or full-text level screening, studies tagged based on PECO eligibility were further categorized based on features such as evidence type (human, animal, mechanistic, PBPK, etc.), health outcome(s), and/or endpoint measure(s) included in the study, and the specific PFAS (or multiple PFAS) addressed (see Appendix B, Table B-7 for

examples). Literature inventories for PECO-relevant studies were created to develop summary-level, sortable lists that include some basic study design information (e.g., study population, exposure information such as doses administered or biomarkers analyzed, age/lifestage[7] of exposure, endpoints examined, etc.). These working literature inventories are for internal use and facilitate subsequent review of individual studies or sets of studies by topic-specific experts.

Inventories were also created for studies that were tagged as "potentially relevant supplemental material" during screening, including *in vitro* or *in silico* models not addressing genotoxicity, ADME studies, and studies on endpoints or routes of exposure that did not meet the specific PECO criteria, but which may still be relevant to the research question(s). Here, the objective was to create an inventory of studies that can be tracked and further summarized as needed—for example, by model system, key characteristic [e.g., of carcinogens, (Smith et al., 2016)], mechanistic endpoint, or key event—to support analyses of potentially critical mechanistic questions that arise at various stages of the systematic review (see Section 9.2 for a description of the process for determining the specific questions and pertinent mechanistic studies to be analyzed). For example, ADME data and related information are important to the next steps of evaluating the evidence from individual PECO-specific studies, and these data will be reviewed by subject matter experts early in the assessment process. Thus, the comprehensive identification of studies relevant to interpreting the ADME or toxicokinetic characteristics of these PFAS was prioritized (see additional discussion in Section 5, and the specifics of the approach in Section 9.2).

[7]Age/lifestage was considered according to EPA's Guidance on Selecting Age Groups for Monitoring and Assessing Childhood Exposures to Environmental Contaminants and EPA's A Framework for Assessing Health Risk of Environmental Exposures to Children.

5. REFINED EVALUATION PLAN

The primary purpose of this step is to outline any potential or expected refinements to the set of populations, exposures, comparators, and outcomes (PECO)-relevant studies, which would narrow the scope of studies considered for use in evidence synthesis and beyond. This optional step is typically applied to focus an assessment with a very large number of PECO-relevant studies on review of the most informative evidence (e.g., when many studies examine the same health outcome, focusing on toxicity studies including exposures below a specified range, those studies examining more specific or objective measures of toxicity, or those that address lifestage- or exposure duration-specific knowledge on how the health outcome develops). Given the relatively small databases of animal toxicology studies for these five PFAS (see Section 2.3.2), this narrowing is not considered applicable to these data. Thus, for these five PFAS assessments, all relevant health outcomes in the animal toxicology studies meeting PECO criteria will be considered.

In contrast to the animal studies, there are many epidemiology studies. To make the systematic review of the epidemiology literature more pragmatic and efficient and focus the set of studies undergoing study evaluation a systematic map of the available evidence was developed after literature screening. The PECO criteria described in Section 3.2 were intentionally broad and inclusive and are well suited for application to systematic mapping. Using the literature inventory, one epidemiologist per outcome reviewed the available evidence and summarized at a high level the direction and consistency of observed associations. In the systematic map (Table 9), the summary of available evidence for each outcome is shaded based on the consistency of direction of association in the studies. These summaries do not account for study risk of bias, with the exception of easily identified critical deficiencies that would make a study or set of studies uninformative (e.g., considerable concern for exposure measurement, confounding, or reverse causation).

Based on the systematic map, outcomes were classified into one of three different tiers of further review based on the likely impact of the outcome on hazard identification and dose-response analyses: (1) systematic review with formal study evaluation (see Section 6.2) with at least two reviewers and evidence synthesis; (2) systematic review with formal study evaluation with one reviewer and evidence synthesis; or (3) systematic map (SM) only - no study evaluation or synthesis of the evidence, although the available database might be mentioned in the assessments to inform data gaps). The depth of review for the first two tiers is equivalent other than the number of reviewers for study evaluation, though the syntheses for the second tier will typically be more succinct due to weaker available evidence.

The determination for review tier was based on the following aspects: (1) consistency of direction of available evidence; (2) consideration of null results in the context of study sensitivity (e.g., if all studies for an outcome reported no association, but also had poor sensitivity due to, for

example, PFAS exposure levels being below or near the limit of detection [LOD], SM only would be used due to an assumed inability to draw any conclusion with confidence); (3) identification of critical deficiencies that would make a study or set of studies uninformative as described above; and (4) consideration of the amount of available animal evidence addressing the human health outcome and the likelihood that a review of the available epidemiology evidence would inform hazard or dose-response conclusions.

In general, an outcome with only uninformative (i.e., having critical deficiencies) or null studies with low sensitivity (shaded light blue in the systematic map) would receive no further review. The systematic review with one reviewer category was used for outcomes in which the available evidence was sparse or the reported results were inconsistent (shaded yellow in the systematic map), but for which there was still a need to conduct a review to provide an informed summary of the available data to support other parts of the assessment (e.g., outcomes where a detailed animal evidence synthesis was expected based on the literature survey in Section 2.3.2; identification of potential research needs or important uncertainties). Generally, outcomes with some or more consistency (shaded light or dark pink in the systematic map) were classified as systematic review with two reviewers, but in a few cases, outcomes with some consistency (shaded light pink) were classified as requiring one reviewer if the complexity appeared to be low (e.g., a small number of studies). In addition, there were logistical reasons that some outcomes were upgraded to two reviewers (e.g., training staff, NTP participation, the same studies used in another outcome receiving two reviews). Outcomes classified as one reviewer can be upgraded if additional complexity is identified during the review.

This approach of tiered reviews is consistent with recommendations from the National Academies of Science encouraging the U.S. Environmental Protection Agency (EPA) to explore ways to make systematic review more feasible, including conducting a "rapid review in which components of the systematic review process are simplified or omitted (e.g., the need for two independent reviewers)" (NAS, 2014).

Table 9. Systematic map of epidemiology outcomes

Health effect/ Outcome	Number of publications[a]					Summary of available evidence	Study evaluation approach (# reviewers)
	PFNA	PFHxS	PFDA	PFHxA	PFBA		
MALE REPRODUCTIVE EFFECTS							
Semen parameters/ sperm DNA damage	5	5	2	1	1	Lower semen parameters in one study for PFNA, PFDA, and PFHxS (ns)	1
Repro hormones	7	7	4	1	0	Mix of higher and lower levels of estradiol and testosterone for PFNA AND PFHxS. No association for PFDA	2[b]
Anogenital distance	1	1	1	0	0	Smaller AGD in boys for PFHxS (significant), but not PFNA or PFDA.	1
Penile width	1	1	1	0	0	No association in single study with low sensitivity	SM
FEMALE REPRODUCTIVE EFFECTS							
Repro hormones	6	6	4	1	0	Mix of higher and lower levels of estradiol and testosterone for PFNA AND PFHxS. No association for PFDA	2[b]
Fecundity	6	6	3	0	0	Three studies observed some positive association with PFNA (one significant), one with PFHxS, but not consistently in subpopulations. One study with inverse association with PFNA, PFHxS, and PFDA.	1
Pubertal timing	1	1	0	0	0	No association in single study with low sensitivity	SM
Menstrual cycle characteristics	2	1	2	0	0	One study reported significant association with irregular cycle; concern for reverse causality. No association reported with PFDA	1
Endometriosis	3	3	2	0	0	Two studies reported association with endometriosis for PFNA (significant) and PFHxS (ns), one study with association for PFDA (ns); concern for reverse causality	1
Menopause	1	1	0	0	0	Earlier age at menopause in one study. High potential for reverse causality	SM
Anogenital distance	1	1	1	0	0	Smaller AGD in girls in single study (significant)	1
DEVELOPMENTAL EFFECTS							
Birth size/fetal growth restriction	17	20	14	0	1	Majority of PFAS studies showed some evidence of birth weight deficits either in the overall population, or among male or female neonates, with statistical significance in some.	2
Preterm birth/ gestational duration	7	8	4	0	1	Shorter gestational duration in 2 studies for PFNA, PFHxS, and PFDA.	2[c]
Postnatal growth	4	5	3	0	0	Inverse association with height and/or weight in 2 studies for PFNA, PFHxS, and PFDA (some significant), but not consistent for all measures. Positive association in 2 studies for PFHxS and 1 study for PFNA (ns).	1
Spontaneous abortion	2	2	2	0	0	Positive association observed in one study, inverse (PFNA and PFDA) or no (PFHxS) association in one study	1
Sex ratio	1	0	1	0	0	No association in single study with low sensitivity	SM

ENDOCRINE EFFECTS							
Thyroid hormones and disease	22	20	10	2	1	Majority of studies reported no association, mix of positive and inverse associations in remaining studies	1
IMMUNE EFFECTS							
Asthma	11	11	8	3	0	Higher asthma in two studies for PFNA and one study for PFHxS (significant)	2[d]
Allergy	6	6	3	0	0	No association in several studies with low sensitivity	2[d]
Antibody response	6	6	4	0	0	All studies of diphtheria and tetanus vaccination reported lower antibody response (some significant) for at least some exposure and outcome measure timing combinations.	2
Infections	4	4	2	0	0	Higher rate of infections observed in two studies for PFNA and PFHxS (significant). No association for PFDA.	2
Atopic dermatitis	6	5	6	0	0	Higher atopic dermatitis in one study for PFNA and two studies for PFHxS (significant)	2[d]
HEPATIC EFFECTS							
Liver enzymes	6	5	1	1	0	Higher liver enzyme levels in two studies for PFNA and PFHxS (significant in one study), lower levels in one study for PFNA (significant). No association for PFDA.	1
Albumin	5	5	1	1	0	No association in studies with low sensitivity; concerns for reverse causality	SM
Liver disease	1	1	1	0	0	Single study evaluated as uninformative	SM
URINARY/RENAL EFFECTS							
Renal function tests	7	8	2	1	0	Associations with impaired renal function with PFNA (3 studies, 2 significant), PFHxS (3 studies, 3 significant), and PFDA (1 study, ns). High potential for reverse causality.	1
CARDIOVASCULAR EFFECTS							
Blood pressure	7	6	4	1	1	Higher odds of hypertension in two studies for PFNA (significant). No association with PFHxS or PFDA.	1
Serum lipids	20	18	9	3	1	Majority of studies report no association; higher serum lipid levels in small number of studies	1
Atherosclerosis	3	2	2	0	0	Higher atherosclerosis in one study for PFNA and PFDA (ns). No association for PFHxS.	1
Coronary heart disease	2	2	2	0	0	No association in two studies with low sensitivity	SM
Ventricular geometry	1	1	1	0	0	Changes in ventricular geometry in single study for PFNA (significant) and PFDA (ns). No association for PFHxS.	1
NERVOUS SYSTEM EFFECTS							
Neurodevelopment	17	15	10	0	0	Small number of studies per specific outcome (e.g., cognition, motor, attention, autism spectrum disorder, cerebral palsy). Mix of positive, inverse, and no associations observed.	1
Adult neurologic effects	3	3	2	0	0	Single study per effect (memory, depression, sleep), with latter two evaluated as uninformative	SM

METABOLIC EFFECTS							
Diabetes	7	6	2	0	0	Higher odds of diabetes in two studies for PFNA and PFHxS (ns) and one study for PFDA (significant). Lower odds in one study for PFHxS (significant) and PFDA (ns).	1
Gestational diabetes	3	3	2	0	0	Two studies report higher odds of gestational diabetes (ns) for PFHxS. No association for PFNA and PFDA.	1
Insulin resistance	12	12	4	1	0	Higher insulin resistance in two studies for PFNA (ns) and PFHxS (one significant). No association for PFDA.	1
Adiposity	10	9	3	0	0	For PFNA and PFDA, most studies report higher adiposity in one outcome measure (ns), but not consistently across measures. No association with adiposity for PFHxS. One study for PFNA, PFDA, and PFHxS reported higher weight gain (significant).	2
Metabolic syndrome	3	3	1	0	0	One study evaluated as uninformative; no association in other studies with low sensitivity	SM
OTHER EFFECTS							
Cancer	6	7	4	0	0	Two studies were evaluated as uninformative. One study available per cancer type with the exception of breast cancer. Higher odds of breast cancer in one study and lower odds in one study (both significant for PFHxS and ns for PFNA).	1
Hematologic effects	1	1	1	1	0	No association in single study with low sensitivity	SM
Mortality	1	1	0	0	0	No association in single study with low sensitivity	SM

[a] Number of publications does not account for multiple publications of the same study
[b] Number of reviewers for study evaluation increased to 2 due to staff training on evaluation of hormonal measures with this approach (see Section 6.2)
[c] Number of reviewers for study evaluation increased to 2 due to the same studies being used for birth size
[d] Number of reviewers for study evaluation increased to 2 because systematic review for immune outcomes was performed by NTP.
SM = systematic map only; ns = not statistically significant; BMI = body mass index; DNA = deoxyribonucleic acid; PFBA = perfluorobutanoic acid; PFDA = perfluorodecanoic acid; PFHxA = perfluorohexanoic acid; PFHxS = perfluorohexanesulfonate; PFNA = perfluorodecanoic acid.

Legend for table shading

Number of publications							
	0		1-4		5-9		10+

Summary of available evidence	
	No association in set of studies or the direction of the association(s) observed is not considered to be toxicologically relevant
	Mix of positive, inverse, and no association results in set of studies, with less than 1/2 of studies in uniform direction. OR 1/2 of studies report association in uniform direction but none are statistically significant OR only one study is available, and results are in the direction of a detrimental effect but are not statistically significant OR >1/2 of studies report association in uniform direction but there is considerable uncertainty due to potential for reverse causality
	>1/2 of studies report association in uniform direction but none are statistically significant OR 1/2-2/3 of studies report association in uniform direction and at least one is statistically significant OR only one study is available and reports statistically significant association
	>2/3 of studies report an association in uniform direction and at least one is statistically significant

To promote consistency in evaluation and presentation across assessments, preliminary decisions were made regarding the grouping of related endpoints for outcome-specific study evaluations and discussion in the evidence synthesis. This helps implement the study evaluation criteria (see Section 6) because those evaluations are outcome and analysis specific. Preliminary decisions for grouping of endpoints from animal toxicology studies for discussion within each assessed human health effect category are described in Table 10. Parallel groupings for outcomes assessed in the available epidemiology studies are captured in Table 9. These groupings are meant to serve as a starting place for consistency in presentation and analysis across studies and assessments, although assessment-specific deviations are possible (e.g., depending on the assessment-specific database of endpoints in the available studies or PFAS-specific understanding of mechanistic relationships across outcomes).

Table 10. Animal Endpoint Grouping Categories

Relevant human health effect category[a]	Examples of animal endpoints included	Notes
General toxicity	• Body weight (not maternal or pup weights, or weights after developmental-only exposure) • Mortality, survival, or LD_{50}s • Growth curve • Clinical observations (non-behavioral)	• Clinical chemistry endpoints are under Hepatic or Hematologic • Maternal or pup body-weight endpoints are under Developmental • Pathology (including gross lesions) is organ specific
Hepatic (toxicity)[b]	• Liver weight and histopathology • Serum or tissue liver enzymes (e.g., ALT and AST from clinical chemistry) • Other liver tissue enzyme activity (e.g., catalase) or protein/DNA content • Other liver tissue biochemical markers (e.g., albumin; glycogen; glucose) • Liver-specific serum biochemistry (e.g., albumin; albumin/globulin) • Liver tissue lipids: triglycerides, cholesterol • Serum lipids (Note: also informative to cardiovascular[d])	• Biochemical markers such as albumin or glucose are under Hematological • Liver tissue cytokines are under Immune • Serum glucose is under Metabolic
Cardiovascular (toxicity)[b,c]	• Heart weight and histopathology • Serum lipids (note: also informative to Hepatic) • Blood pressure • Blood measures of cardiovascular risk (e.g., C-reactive protein)	• Other blood measures are under Hepatic, Immune, or Hematologic
Hematologic (effects)[b,c]	• Red blood cells • Blood hematocrit or hemoglobin • Corpuscular volume • Blood platelets or reticulocytes • Blood biochemical measures (e.g., sodium, calcium, phosphorus)	• White blood cell count and globulin are under Immune • Serum lipids are under Cardiovascular • Serum liver markers are under Hepatic
Immune (effects)[b]	• Host resistance	• Red blood cells are under Hematological

Relevant human health effect category[a]	Examples of animal endpoints included	Notes
	• Allergic, autoimmune or infectious disease • Hypersensitivity • General immune assays (e.g., white blood cell counts, immunological factors or cytokines in blood, lymphocyte phenotyping or proliferation) • Any measure in lymphoid tissues (weight; histopathology; cell counts; etc.) • Immune cell counts or immune-specific cytokines in non-lymphoid tissues • Other immune functional assays (e.g., antibody production, natural killer cell function, DTH, MLR, CTL, phagocytosis or bacterial killing by monocytes) • Immune responses in the respiratory system • Stress-related factors in blood (e.g., glucocorticoids or other adrenal markers)	• Non-immune measures of pulmonary function are under Respiratory
Urinary (toxicity)[b]	• Kidney weight and histopathology • Urinary measures (e.g., protein, volume, pH, specific gravity)	
Nervous system (effects)[b,c]	• Brain weight • Behavioral measures (including FOB and cage-side observations) • Nervous system histopathology	
Endocrine (effects)[b]	• Thyroid weight and histopathology • Hormonal measures in any tissue or blood (non-reproductive)	• Reproductive hormones are under Reproductive

Relevant human health effect category[a]	Examples of animal endpoints included	Notes
Metabolic (effects)[b,c]	• Free fatty acids • Serum glucose or insulin, or other measures related to diabetes • Pancreatic effects relevant to diabetes • Induced-obesity or BMI • Any of the above endpoints after developmental exposure will be primarily discussed in this health effect category, and then referenced under developmental effects	
Reproductive (toxicity)[b] Note: Evidence synthesis and evidence integration conclusions in assessments are developed separately for male and female reproductive effects (toxicity)	• Reproductive organ weight and histopathology • Markers of sexual differentiation or maturation (e.g., preputial separation in males; vaginal opening or estrous cycling in females) • Mating parameters (e.g., success, mount latency) • Reproductive hormones	• Birth parameters (e.g., litter size; resorptions, implantations, viability) are under Developmental • If data indicate altered birth parameters are likely attributable to female fertility, these data may be discussed under Female Reproductive
Developmental (effects)[b] Note: Evidence synthesis of these endpoints in the assessments is termed "offspring growth and early development," but evidence integration conclusions will be drawn on the broader category of "developmental effects" (which also considers organ/system-specific effects after exposure during development)	• Dam health (e.g., weight gain, food consumption) • Pup viability/survival or other birth parameters (e.g., number of pups per litter) • Pup weight or growth (includes measures into adulthood after developmental-only exposure) • Developmental landmarks (eye opening, etc., but not including markers for other organ/system-specific toxicities)	• Histopathology and markers of development specific to other systems are organ/system-specific (e.g., vaginal opening is under Female Reproductive; tests of sensory maturation are under Nervous System)

Relevant human health effect category[a]	Examples of animal endpoints included	Notes
Carcinogenicity[b]	• Tumors • Precancerous lesions (e.g., dysplasia)	

ALT = alanine aminotransferase; AST = aspartate transaminase; BMI = body mass index; CTL = cytotoxic T lymphocyte; DNA = deoxyribonucleic acid; DTH = delayed-type hypersensitivity; FOB = functional operational battery; LD_{50} = median lethal dose; MLR = mixed leukocyte reaction.

[a]Given the paucity of available studies and the absence of exceptional new evidence, information on gastrointestinal effects, musculoskeletal effects, ocular effects, and respiratory effects will not be formally evaluated in these assessments, although short summaries of the evidence may be included for context, and new literature relating to these outcomes will be monitored during literature search updates for potential inclusion.

[b]Any of the health effect-relevant endpoints observed after developmental exposure will be discussed primarily in the health effect category indicated, and then referenced under developmental effects.

[c]The primary focus of these assessments will be on developmental effects; endocrine effects; hepatic effects, including lipid and lipoprotein measures; immune effects; reproductive effects in males or females; urinary effects; general toxicity; and carcinogenicity. Data on cardiovascular effects, hematological effects, metabolic effects including diabetes, and nervous system effects will be summarized when available. These summaries may be developed in association with one of the health effects noted above, as a separate formal evaluation of hazard, or as part of a qualitative summary on "other effects," depending on the assessment-specific data.

[d]Some outcomes are relevant to multiple health effects. These outcomes may be categorized under only a single health effect in Table 10 for clarity. However, in the assessments, such outcome data would be discussed in the first relevant health effect synthesis (syntheses will generally follow the pattern of most-to-least available evidence) and then this synthesis will be cited in the syntheses of other relevant health effects. The evidence (for or against an effect) will contribute to evidence integration decisions for all relevant health effects.

Assessment-specific refinements to the evaluation plan described in later sections of this protocol may be justified after review of the key areas of scientific complexity outlined in Section 2.4. Although not expected based on the relatively small database of studies for these PFAS, one such refinement includes the potential prioritization of studies testing specific (lower) exposure levels, exposure lifestages, routes of exposure, or toxic metabolites as identified based on conclusions made regarding the ADME properties of these PFAS. As noted in Section 2.4, consideration of the available ADME data for these five PFAS will be prioritized (see additional discussion in Section 9.2). This will serve multiple purposes, including updating the data in Table 7 on serum half-lives across species and sexes. Notably, it is not expected that there will be enough data to examine lifestage-specific differences in ADME (including metabolic pathways for toxification or detoxification) that might inform evidence evaluation and synthesis decisions. (This is distinct from lifestage-specific differences in *exposure*, e.g., due to the higher intake of food per kg body weight [BW] of young children or ingestion of dust.) However, a few anticipatory refinements will be applied to study evaluations based on the preliminary data presented in Table 7. Specifically, given the apparent sex-specific differences in PFAS half-life in rats and mice (note: toxicology studies in nonhuman primates were not identified for these PFAS), examining and reporting data for both sexes will be reviewed as a potential source of study insensitivity during

study evaluation (see Section 6.3), particularly for PFAS that seem to vary largely for this parameter (e.g., PFHxS; PFNA). These half-life data will also be considered when evaluating the experiment-specific sensitivity of the frequency of exposures and the timing of endpoint testing after exposure in experimental animals (see Section 6.3), as well as the potential for using exposure biomarkers in exposed animals and humans (see Sections 6.2 and 6.3). The apparent ADME differences across species will be a critical consideration in these assessments. This consideration will be applied to evidence synthesis and integration decisions (e.g., exploring ADME differences between rats and mice as a potential explanation if there are differences in sensitivity in outcome-specific responses; see Sections 9 and 10), as well as in extrapolating dosimetry (i.e., exposure levels and duration) from experiments in animal models to quantitative estimates relevant to humans, possibly including application within existing pharmacokinetic (PK) and PBPK models (see additional discussion in Sections 6.4 and 11.2).

Lastly, based on the key areas of scientific complexity outlined in Section 2.4, some of the analyses performed in support of these assessments may need to consider a broader array of studies than those available for these five PFAS. One example includes the need to consider the human relevance of certain outcomes observed in animals, including the role of receptors such as PPARα (see additional discussion in Section 9.2). In addition, it is possible that additional literature identification and evaluation strategies will be developed to address the other key areas of scientific complexity outlined in Section 2.4, or other assessment-specific issues that arise during review. Any such approaches will be documented in updates to the PFAS-specific assessment protocol(s) (i.e., as assessment-specific updates to this document included as appendix materials for each of the five PFAS assessments).

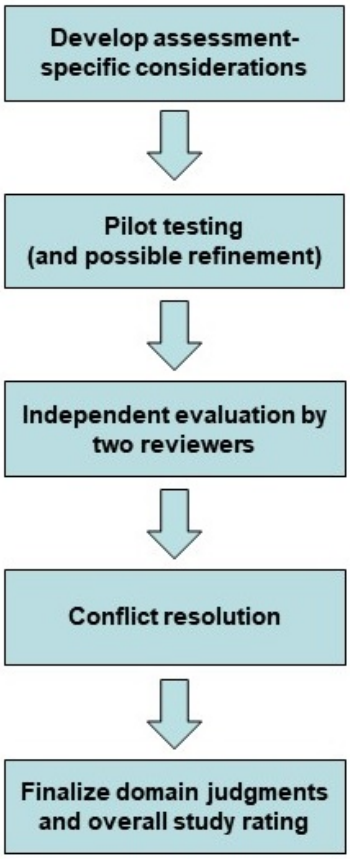
6. STUDY EVALUATION (REPORTING, RISK OF BIAS, AND SENSITIVITY) STRATEGY

The general approach for evaluating PECO-relevant primary health effect studies is described in Section 6.1 and is the same for epidemiology and animal toxicology experiments, but the specifics of applying the approach differ; thus, they are described separately for epidemiology and animal toxicology studies in Sections 6.2 and 6.3, respectively. No controlled human exposure studies for these PFAS were identified (see Section 4). Although they have not yet been formally identified by the systematic literature searches (this is expected during the next literature search update), PBPK modeling studies were recently identified for PFHxS (Kim et al., 2018) and for PFDA and PFNA (Kim et al., 2019). In addition, a two-compartment PK model for gestational and lactational transfer of PFHxS in humans has been described by Verner et al. (2016). The specific approach for reviewing these studies is described in Section 6.4. Different approaches are used to evaluate mechanistic studies (see Sections 6.5 and 9.2).

6.1. STUDY EVALUATION OVERVIEW FOR HEALTH EFFECT STUDIES

Key concerns for the review of epidemiology and animal toxicology studies are potential bias (factors that affect the magnitude or direction of an effect in either direction) and insensitivity (factors that limit the ability of a study to detect a true effect; low sensitivity is a bias towards the null when an effect exists). The study evaluations are aimed at discerning the expected magnitude of any identified limitations (focusing on limitations that could substantively change a result), considering also the expected direction of the bias. The study evaluation approach is designed to address a range of study designs, health effects, and chemicals. The general approach for reaching an overall judgment for the study (or a specific analysis within a study) regarding confidence in the reliability of the results is illustrated in Figure 5.

(a) Study evaluation process



(b) Individual evaluation domains

Animal	Epidemiology
Selection and performance • Allocation • Observational bias/blinding	Participant selection
Confounding/variable control	Confounding
Selective reporting and attrition	Selective reporting
Exposure methods sensitivity • Chemical administration and characterization • Exposure timing, frequency, and duration	Exposure measurement
Outcome measures and results display • Endpoint sensitivity and specificity • Results presentation	Outcome ascertainment Analysis
Reporting quality	Other sensitivity

Domain judgments

Judgment	Interpretation
Good	Appropriate study conduct relating to the domain and minor deficiencies not expected to influence results.
Adequate	A study that may have some limitations relating to the domain, but they are not likely to be severe or to have a notable impact on results.
Deficient	Identified biases or deficiencies interpreted as likely to have had a notable impact on the results or prevent reliable interpretation of study findings.
Critically Deficient	A serious flaw identified that makes the observed effect(s) uninterpretable. Studies with a critical deficiency will almost always be considered "uninformative" overall.

Overall study rating for an outcome

Rating	Interpretation
High	No notable deficiencies or concerns identified; potential for bias unlikely or minimal; sensitive methodology.
Medium	Possible deficiencies or concerns noted, but resulting bias or lack of sensitivity is unlikely to be of a notable degree.
Low	Deficiencies or concerns were noted, and the potential for substantive bias or inadequate sensitivity could have a significant impact on the study results or their interpretation.
Uninformative	Serious flaw(s) makes study results unusable for hazard identification or dose response.

Figure 5. Overview of Integrated Risk Information System (IRIS) study evaluation process. (a) An overview of the general evaluation process [Note: see Section 5 for deviations from independent evaluation by two reviewers for some health outcomes in epidemiology studies]. (b) The evaluation domains and definitions for ratings (i.e., domain and overall judgments, performed on an outcome-specific basis).

1

With the exceptions noted in the refined evaluation plan for select outcomes reported in epidemiology studies (see Section 5), at least two reviewers will independently evaluate health effect studies to identify characteristics that bear on the informativeness (i.e., validity and sensitivity) of the results. The independent reviewers will use the structured platform for study evaluation housed within the Environmental Protection Agency's (EPA's) version of the Health Assessment Workplace Collaboration (HAWC)[8] to record separate judgements for each domain and the overall study for each outcome, to reach consensus between reviewers, and when necessary, resolve differences by discussion between the reviewers or consultation with additional independent reviewers. For some domains, additional chemical- or outcome-specific knowledge will be applied to evaluating the experimental design and methodology, as described below.

In general, considerations for reviewing a study with regard to its conduct for specific health outcomes is based on the use of existing guidance documents when available, including EPA guidance for carcinogenicity, neurotoxicity, reproductive toxicity, and developmental toxicity (U.S. EPA, 2005a, 1998, 1996a, 1991a). For some aspects of the study evaluations (e.g., review of exposure assessment in epidemiology studies), additional considerations are developed in consultation with topic-specific technical experts. To calibrate the assessment-specific considerations, the study evaluations will include a pilot phase to assess and refine the evaluation process. Additionally, as reviewers examine a group of studies, additional chemical-specific knowledge or methodologic concerns may emerge and a second pass of all pertinent studies may become necessary. Refinements to the study evaluation process made during the pilot phase and subsequent implementation across all relevant studies will be acknowledged as updates to the protocol.

Authors may be queried to obtain critical information, particularly that involving missing reporting quality information or other data (e.g., information on variability) or additional analyses that could address potential study limitations. The decision on whether to seek missing information includes consideration of what additional information would be useful, specifically with respect to any information that could result in a reevaluation of the overall study confidence for an outcome. Outreach to study authors will be documented and considered unsuccessful if researchers do not respond to an email or phone request within one month of the attempt to contact. Only information or data that can be made publicly available (e.g., within HAWC or Health and Environmental Research Online [HERO]) will be considered.

[8]HAWC is a free and open source software application that provides a modular, web-based interface to help develop human health assessments of chemicals: https://hawcproject.org/portal/. Standard operating procedures provided to the reviewers to facilitate consistent and relevant documentation of their judgments using the HAWC software can be found as attachments embedded within the online tool (https://hawcprd.epa.gov/assessment/100000039/).

When evaluating studies[9] that examine more than one outcome, the evaluation process will be performed separately for each outcome, because the utility of a study can vary for different outcomes. If a study examines multiple endpoints for the same outcome,[10] evaluations may be performed at a more granular level if appropriate, but these measures may still be grouped for evidence synthesis.

During review, for each evaluation domain reviewers will reach a consensus judgment of *good*, *adequate*, *deficient*, *not reported*, or *critically deficient*. If a consensus is not reached, a third reviewer will perform conflict resolution. It is important to stress that these evaluations are performed in the context of the study's utility for identifying individual hazards. While limitations specific to the usability of the study for dose-response analysis are useful to note to inform those later decisions, they do not contribute to the study confidence classifications.

These four categories are applied to each evaluation domain for each study as follows:

- *Good* represents a judgment that the study was conducted appropriately in relation to the evaluation domain, and any minor deficiencies that are noted would not be expected to influence the study results.

- *Adequat*e indicates a judgment that there may be methodological limitations relating to the evaluation domain, but that those limitations are not likely to be severe or to have a notable impact on the results.

- *Deficient* denotes identified biases or deficiencies that are interpreted as likely to have had a notable impact on the results or that prevent interpretation of the study findings.

- *Not reported* indicates that the information necessary to evaluate the domain question was not available in the study. Generally, this term carries the same functional interpretation as *deficient* for the purposes of the study confidence classification (described below). Depending on the number of unreported items and severity of other limitations identified in the study, it may or may not be worth reaching out to the study authors for this information (see discussion above).

- *Critically deficient* reflects a judgment that the study conduct relating to the evaluation domain question introduced a serious flaw that is interpreted to be the primary driver of any observed effect(s) or makes the study uninterpretable. Studies with a determination of *critically deficient* in an evaluation domain will not be used for hazard identification or dose-response analysis, but they may be used to highlight possible research gaps.

[9]Note: "study" is used instead of a more accurate term (e.g., "experiment") throughout these sections owing to an established familiarity within the field for discussing a study's risk of bias or sensitivity, etc. However, all evaluations discussed herein are explicitly conducted at the level of an individual outcome within a population or cohort of humans or animals exposed in a similar manner (e.g., unexposed or exposed at comparable lifestages and for the same duration of exposure), or to a sample of the study population within a study.

[10]Note: "outcome" will be used throughout these methods; the same methods also apply to an endpoint assessed separately within a larger outcome. "Endpoint" refers to a more granular measurement (e.g., for the outcome of liver histopathology, different endpoints might include necrosis and cellular hypertrophy).

Once the evaluation domains have been rated, the identified strengths and limitations will be considered as a whole to reach a study confidence classification of *high*, *medium*, or *low* confidence, or *uninformative* for a specific health outcome. This classification is based on the reviewer judgments across the evaluation domains and considers the likely impact that inadequate reporting or the noted deficiencies in bias and sensitivity have on the outcome-specific results. The classifications, which reflect a consensus judgment between reviewers, are defined as follows:

- *High* confidence: No notable deficiencies or concerns were identified; the potential for bias is unlikely or minimal, and the study used sensitive methodology. *High*-confidence studies generally reflect judgments of *good* across all or most evaluation domains.

- *Medium* confidence: Possible deficiencies or concerns were noted, but the limitations are unlikely to be of a notable degree. Generally, *medium*-confidence studies include *adequate* or *good* judgments across most domains, with the impact of any identified limitation not being judged as severe.

- *Low* confidence: Deficiencies or concerns are noted, and the potential for bias or inadequate sensitivity could have a significant impact on the study results or their interpretation. Typically, *low*-confidence studies have a *deficient* evaluation for one or more domains, although some *medium*-confidence studies may have a *deficient* rating in domain(s) considered to have less influence on the magnitude or direction of the outcome-specific results). *Low*-confidence results are given less weight compared to *high*- or m*edium*-confidence results during evidence synthesis and integration (see Section 10.1, Table 20and Table 21), and are generally not used on their own for hazard identification or dose-response analyses unless they are the only studies available or they inform data gaps unexamined in the *high*- or *medium*-confidence studies. Studies rated as *medium*- or *low*-confidence only because of sensitivity concerns about bias towards the null will be asterisked or otherwise noted because they may require additional consideration during evidence synthesis. Effects observed in studies biased toward the null may actually increase confidence in the results, assuming the study is otherwise well conducted (see Section 9).

- *Uninformative*: Serious flaw(s) make the study results unusable for hazard identification. Studies with *critically deficient* judgements in any evaluation domain are almost always classified as *uninformative* (see explanation above). Studies with multiple *deficient* judgments across domains may also be considered *uninformative*. *Uninformative* studies will not be considered further in the synthesis and integration of evidence, except perhaps to highlight possible research gaps.

As previously noted, study evaluation determinations reached by each reviewer and the consensus judgment between reviewers will be recorded in HAWC. Final study evaluations housed in HAWC will be made available when the draft is publicly released. The study confidence classifications and their rationales will be carried forward and considered as part of evidence synthesis (see Section 9) to help interpret the results across studies.

6.2. EPIDEMIOLOGY STUDY EVALUATION

Evaluation of epidemiology studies of health effects to assess risk of bias and study sensitivity will be conducted for the following domains: exposure measurement, outcome ascertainment, participant selection, potential confounding, analysis, study sensitivity, and selective reporting. Bias can result in false positives and negatives (i.e., Types I and II errors), while study sensitivity is typically concerned with identifying the latter.

The principles and framework used for evaluating epidemiology studies are based on the Cochrane Risk of Bias in Nonrandomized Studies of Interventions [ROBINS-I; (Sterne et al., 2016)] but modified to address environmental and occupational exposures. The underlying philosophy of ROBINS-I is to describe attributes of an "ideal" study with respect to each of the evaluation domains (e.g., exposure measurement, outcome classification, etc.). Core and prompting questions are used to collect information to guide evaluation of each domain.

Core and prompting questions for each domain are presented in Table 11. Core questions represent key concepts, while the prompting questions help the reviewer focus on relevant details under each key domain. Table 11 also includes criteria that apply to all exposures and outcomes. PFAS-specific criteria are described in Section 6.2.1. As mentioned in Section 6.1, any additions to or refinements of the criteria will be documented as updates to the protocol.

Table 11. Questions and criteria for evaluating each domain in epidemiology studies

Domain and core question	Prompting questions	Follow-up questions	Criteria that apply to most exposures and outcomes
<u>Exposure measurement</u> Does the exposure measure reliably distinguish between levels of exposure in a time window considered most relevant for a causal effect with respect to the development of the outcome?	For all: • Does the exposure measure capture the variability in exposure among the participants, considering intensity, frequency, and duration of exposure? • Does the exposure measure reflect a relevant time window? If not, can the relationship between measures in this time and the relevant time window be estimated reliably? • Was the exposure measurement likely to be affected by a knowledge of the outcome? • Was the exposure measurement likely to be affected by the presence of the outcome (i.e., reverse causality)? For case-control studies of occupational exposures: • Is exposure based on a comprehensive job history describing tasks, setting, time period, and use of specific materials? For biomarkers of exposure, general population: • Is a standard assay used? What are the intra- and inter-assay coefficients of variation? Is the assay likely to be affected by contamination? Are values less than	Is the degree of exposure misclassification likely to vary by exposure level? If the correlation between exposure measurements is moderate, is there an adequate statistical approach to ameliorate variability in measurements? If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?	**Good** • Valid exposure assessment methods used, which represent the etiologically relevant time period of interest. • Exposure misclassification is expected to be minimal. **Adequate** • Valid exposure assessment methods used, which represent the etiologically relevant time period of interest. • Exposure misclassification may exist but is not expected to greatly change the effect estimate. **Deficient** • Valid exposure assessment methods used, which represent the etiologically relevant time period of interest. Specific knowledge about the exposure and outcome raise concerns about reverse causality, but there is uncertainty whether it is influencing the effect estimate. • Exposed groups are expected to contain a notable proportion of unexposed or minimally exposed individuals, the method did not capture important temporal or spatial variation, or there is other evidence of exposure misclassification that would be expected to notably change the effect estimate.

Domain and core question	Prompting questions	Follow-up questions	Criteria that apply to most exposures and outcomes
	the limit of detection dealt with adequately? • What exposure time period is reflected by the biomarker? If the half-life is short, what is the correlation between serial measurements of exposure?		**Critically deficient** • Exposure measurement does not characterize the etiologically relevant time period of exposure or is not valid. • There is evidence that reverse causality is very likely to account for the observed association. • Exposure measurement was not independent of outcome status.
<u>Outcome ascertainment</u> Does the outcome measure reliably distinguish the presence or absence (or degree of severity) of the outcome?	For all: • Is outcome ascertainment likely to be affected by knowledge of, or presence of, exposure (e.g., consider access to health care, if based on self-reported history of diagnosis)? For case-control studies: • Is the comparison group without the outcome (e.g., controls in a case-control study) based on objective criteria with little or no likelihood of inclusion of people with the disease? For mortality measures: • How well does cause-of-death data reflect occurrence of the disease in an individual? How well do mortality data reflect incidence of the disease? For diagnosis of disease measures:	Is there a concern that any outcome misclassification is non-differential, differential, or both? What is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?	**Good** • High certainty in the outcome definition (i.e., specificity and sensitivity), minimal concerns with respect to misclassification. • Assessment instrument was validated in a population comparable to the one from which the study group was selected. **Adequate** • Moderate confidence that outcome definition was specific and sensitive, some uncertainty with respect to misclassification but not expected to greatly change the effect estimate. • Assessment instrument was validated but not necessarily in a population comparable to the study group. **Deficient** • Outcome definition was not specific or sensitive. • Uncertainty regarding validity of assessment instrument.

Domain and core question	Prompting questions	Follow-up questions	Criteria that apply to most exposures and outcomes
	• Is the diagnosis based on standard clinical criteria? If it is based on self-report of the diagnosis, what is the validity of this measure? For laboratory-based measures (e.g., hormone levels): • Is a standard assay used? Does the assay have an acceptable level of inter-assay variability? Is the sensitivity of the assay appropriate for the outcome measure in this study population?		**Critically deficient** • Invalid/insensitive marker of outcome. • Outcome ascertainment is very likely to be affected by knowledge of, or presence of, exposure. Note: Lack of blinding should not be automatically construed to be *critically deficient*.
<u>Participant selection</u> Is there evidence that selection into or out of the study (or analysis sample) was jointly related to exposure and to outcome?	For longitudinal cohort: • Did participants volunteer for the cohort based on knowledge of exposure and/or preclinical disease symptoms? Was entry into the cohort or continuation in the cohort related to exposure and outcome? For occupational cohort: • Did entry into the cohort begin with the start of the exposure? • Was follow-up or outcome assessment incomplete, and if so, was follow-up related to both exposure and outcome status? • Could exposure produce symptoms that would result in a change in work assignment/work status ("healthy worker survivor effect")?	Were differences in participant enrollment and follow-up evaluated to assess bias? If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?	**Good** • Minimal concern for selection bias based on description of recruitment process (e.g., selection of comparison population, population-based random sample selection, recruitment from sampling frame including current and previous employees). • Exclusion and inclusion criteria specified and would not induce bias. • Participation rate is reported at all steps of study (e.g., initial enrollment, follow-up, selection into analysis sample). If rate is not high, there is appropriate rationale for why it is unlikely to be related to exposure (e.g., comparison between participants and nonparticipants or other available information indicates differential selection is not likely). **Adequate** • Enough of a description of the recruitment process to be comfortable that there is no serious risk of bias.

Domain and core question	Prompting questions	Follow-up questions	Criteria that apply to most exposures and outcomes
	For case-control study: • Were controls representative of population and time periods from which cases were drawn? • Are hospital controls selected from a group whose reason for admission is independent of exposure? • Could recruitment strategies, eligibility criteria, or participation rates result in differential participation relating to both disease and exposure? For population-based survey: • Was recruitment based on advertisement to people with knowledge of exposure, outcome, and hypothesis?	Were appropriate analyses performed to address changing exposures over time in relation to symptoms? Is there a comparison of participants and nonparticipants to address whether differential selection is likely?	• Inclusion and exclusion criteria specified and would not induce bias. • Participation rate is incompletely reported but available information indicates participation is unlikely to be related to exposure. **Deficient** • Little information on recruitment process, selection strategy, sampling framework, and/or participation OR aspects of these processes raises the potential for bias (e.g., healthy worker effect, survivor bias). **Critically deficient** • Aspects of the processes for recruitment, selection strategy, sampling framework, or participation result in concern that selection bias is likely to have had a large impact on effect estimates (e.g., convenience sample with no information about recruitment and selection, cases and controls are recruited from different sources with different likelihood of exposure, recruitment materials stated outcome of interest and potential participants are aware of or are concerned about specific exposures).
<u>Confounding</u> Is confounding of the effect of the exposure likely?	Is confounding adequately addressed by considerations in: • Participant selection (matching or restriction)? • Accurate information on potential confounders and statistical adjustment procedures?	If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is	**Good** • Conveys strategy for identifying key confounders. This may include *a priori* biological considerations, published literature, causal diagrams, or statistical analyses, with the recognition that not all "risk factors" are confounders. • Inclusion of potential confounders in statistical models not based solely on statistical significance criteria (e.g., $p < 0.05$ from stepwise regression).

Domain and core question	Prompting questions	Follow-up questions	Criteria that apply to most exposures and outcomes
	• Lack of association between confounder and outcome, or confounder and exposure in the study? • Information from other sources? Is the assessment of confounders based on a thoughtful review of published literature, potential relationships (e.g., as can be gained through directed acyclic graphing), and minimizing potential overcontrol (e.g., inclusion of a variable on the pathway between exposure and outcome)?	enough information)?	• Does not include variables in the models that are likely to be influential colliders or intermediates on the causal pathway. • Key confounders are evaluated appropriately and considered to be unlikely sources of substantial confounding. This often will include: o Presenting the distribution of potential confounders by levels of the exposure of interest and/or the outcomes of interest (with amount of missing data noted); o Consideration that potential confounders were rare among the study population, or were expected to be poorly correlated with exposure of interest; o Consideration of the most relevant functional forms of potential confounders; o Examination of the potential impact of measurement error or missing data on confounder adjustment; or o Presenting a progression of model results with adjustments for different potential confounders, if warranted. **Adequate** • Similar to *good* but may not have included all key confounders, or less detail may be available on the evaluation of confounders (e.g., sub-bullets in *good*). It is possible that residual confounding could explain part of the observed effect, but concern is minimal.

Domain and core question	Prompting questions	Follow-up questions	Criteria that apply to most exposures and outcomes
			Deficient • Does not include variables in the models that have been shown to be influential colliders or intermediates on the causal pathway. • And any of the following: o The potential for bias to explain some of the results is high based on an inability to rule out residual confounding, such as a lack of demonstration that key confounders of the exposure-outcome relationships were considered; o Descriptive information on key confounders (e.g., their relationship relative to the outcomes and exposure levels) are not presented; or o Strategy of evaluating confounding is unclear or is not recommended (e.g., only based on statistical significance criteria or stepwise regression [forward or backward elimination]). **Critically deficient** • Includes variables in the models that are colliders and/or intermediates in the causal pathway, indicating that substantial bias is likely from this adjustment; or • Confounding is likely present and not accounted for, indicating that all the results were most likely due to bias.
<u>Analysis</u> Does the analysis strategy and presentation convey the necessary	• Are missing outcome, exposure, and covariate data recognized, and if necessary, accounted for in the analysis?	If there is a concern about the potential for bias, what is the predicted direction or	**Good** • Use of an optimal characterization of the outcome variable, including presentation of subgroup- or lifestage-specific comparisons (as appropriate for the outcome).

Domain and core question	Prompting questions	Follow-up questions	Criteria that apply to most exposures and outcomes
familiarity with the data and assumptions?	• Does the analysis appropriately consider variable distributions and modeling assumptions? • Does the analysis appropriately consider subgroups or lifestages of interest (e.g., based on variability in exposure level or duration or susceptibility)? • Is an appropriate analysis used for the study design? • Is effect modification considered, based on considerations developed a priori? • Does the study include additional analyses addressing potential biases or limitations (i.e., sensitivity analyses)?	distortion of the bias on the effect estimate (if there is enough information)?	• Quantitative results presented (effect estimates and confidence limits or variability in estimates) (i.e., not presented only as a *p*-value or "significant"/"not significant"). • Descriptive information about outcome and exposure provided (where applicable). • Amount of missing data noted and addressed appropriately (discussion of selection issues—missing at random vs. differential). • Where applicable, for exposure, includes LOD (and percentage below the LOD), and decision to use log transformation. • Includes analyses that address robustness of findings, e.g., examination of exposure-response (explicit consideration of nonlinear possibilities, quadratic, spline, or threshold/ceiling effects included, when feasible); relevant sensitivity analyses; effect modification examined based only on a priori rationale with sufficient numbers. • No deficiencies in analysis evident. Discussion of some details may be absent (e.g., examination of outliers). **Adequate** • Same as 'Good', except: • Descriptive information about exposure provided (where applicable) but may be incomplete; might not have discussed missing data, cut-points, or shape of distribution(s). • Includes analyses that address robustness of findings (examples in 'Good'), but some important analyses are not performed.

Domain and core question	Prompting questions	Follow-up questions	Criteria that apply to most exposures and outcomes
			Deficient • Does not conduct analysis using optimal characterization of the outcome variable. • Descriptive information about exposure levels not provided (where applicable). • Effect estimate and *p*-value presented, without standard error or confidence interval. • Results presented as statistically "significant"/"not significant." **Critically deficient** • Results of analyses of effect modification examined without clear a priori rationale and without providing main/principal effects (e.g., presentation only of statistically significant interactions that were not hypothesis driven). • Analysis methods are not appropriate for design or data of the study.
<u>Selective reporting</u> Is there reason to be concerned about selective reporting?	• Were results provided for all the primary analyses described in the methods section? • Is there appropriate justification for restricting the amount and type of results that are shown? • Are only statistically significant results presented?	If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?	**Good** • The results reported by study authors are consistent with the primary and secondary analyses described in a registered protocol or methods paper. **Adequate** • The authors described their primary (and secondary) analyses in the methods section and results were reported for all primary analyses.

Domain and core question	Prompting questions	Follow-up questions	Criteria that apply to most exposures and outcomes
			Deficient • Concerns were raised based on previous publications, a methods paper, or a registered protocol indicating that analyses were planned or conducted that were not reported, or that hypotheses originally considered to be secondary were represented as primary in the reviewed paper. • Only subgroup analyses were reported, suggesting that results for the entire group were omitted. • Only statistically significant results were reported.
Sensitivity Is there a concern that sensitivity of the study is not adequate to detect an effect?	• Is the exposure range adequate to detect associations and exposure-response relationships? • Was the appropriate population or lifestage included? • Was the length of follow-up adequate? Is the time/age of outcome ascertainment optimal given the interval of exposure and the health outcome? • Are there other aspects related to risk of bias or otherwise that raise concerns about sensitivity?		**Adequate** • The range of exposure levels provides adequate variability to evaluate the associations of relevance. • The population was exposed to levels expected to have an impact on response. • The study population was sensitive to the development of the outcomes of interest (e.g., ages, lifestage, sex). • The timing of outcome ascertainment was appropriate given expected latency for outcome development (i.e., adequate follow-up interval). • The study was adequately powered to observe an effect. • No other concerns raised regarding study sensitivity. **Deficient** • Concerns were raised about the issues described for *adequate* that are expected to notably decrease the sensitivity of the study to detect associations for the outcome.

1

6.2.1. Epidemiology Study Evaluation Criteria Specific to These Five Per- and Polyfluoroalkyl Substances (PFAS)

The exposure criteria described in Table 12 below are modified from the criteria developed by NTP OHAT[11] for their assessment of the association between PFOA and immune effects.

The estimated serum half-lives of PFAS in humans were presented in Table 7 (see Section 2.4.1). In considering temporality concerns, some PFAS (PFHxS, PFDA, and PFNA) are persistent compounds with longer (multiple year) half-lives in humans, so current exposure levels may be indicative of critical exposure windows that were narrow or past exposures that extended beyond the anticipated half-lives. In contrast, other PFAS appear to have half-lives of 1 month or less (PFBA, PFHxA), and current exposure levels may not be indicative of past exposures that extend beyond the anticipated half-lives. Some evidence suggests that the half-lives vary based on sex, parity, interval between pregnancy, reproductive hormones, and gynecological disorders (Lau et al., 2007); therefore, these factors will be considered depending on the population(s), critical windows, and outcomes being examined.

Standard analytical methods of individual PFAS in serum or whole-blood using quantitative techniques such as liquid chromatography-triple quadrupole mass spectrometry are considered to be well-established methods (CDC, 2019a, b; ATSDR, 2018; CDC, 2015; U.S. EPA, 2014a, b; CDC, 2009).

[11]National Toxicology Program (NTP) Report:
https://ntp.niehs.nih.gov/ntp/ohat/pfoa_pfos/pfoa_pfosmonograph_508.pdf
NTP protocol: https://ntp.niehs.nih.gov/ntp/ohat/pfoa_pfos/protocol_201506_508.pdf.

Table 12. Criteria for evaluating exposure measurement in epidemiology studies of per- and polyfluoroalkyl substances (PFAS) and health effects

Rating	Criteria
Good	• Evidence that exposure was consistently assessed using well-established analytical methods that directly measure exposure (e.g., measurement of PFAS in blood, serum, or plasma). **OR** • Exposure was assessed using less established methods (e.g., measurement of PFAS in breast milk) or methods that indirectly measure exposure (e.g., drinking water concentrations and residential location/history, questionnaire or occupational exposure assessment by a certified industrial hygienist) that are validated against well-established direct methods (i.e., inter-methods validation: one method vs. another) in the target population of interest. **And all the following:** • Exposure was assessed in a relevant time-window (i.e., temporality is established, and sufficient latency occurred prior to disease onset) for development of the outcome based on current biological understanding. • There is evidence that sufficient exposure data measurements are above the limit of quantification for the assay. • The laboratory analysis included data on standard quality control measures with demonstrated precision and accuracy.
Adequate	• Exposure was assessed using less established methods or indirect measures that are validated but not in the target population of interest. **OR** • Evidence that exposure was consistently assessed using methods described in 'Good', but there were some concerns about quality control measures or other potential for non-differential misclassification. **And all the following:** • Exposure was assessed in a relevant time-window for development of the outcome • There is evidence that sufficient exposure data measurements are above the limit of quantification for the assay. • The laboratory analysis included some data on standard quality control measures with demonstrated precision and accuracy.
Deficient	**Any of the following:** • Some concern, but no direct evidence, that the exposure was assessed using methods that have not been validated or empirically shown to be consistent with methods that directly measure exposure. • Exposure was assessed in a relevant time window(s) for development of the outcome, but there could be some concern about the potential for bias due to reverse causality[a] between exposure and outcome, yet no direct evidence that it is present.

Rating	Criteria
Critically deficient	**Any of the following:** • Exposure was assessed in a time window that is unknown or not relevant for development of the outcome. This could be due to clear evidence of bias from reverse causality between exposure and outcome, or other concerns such as the lack of temporal ordering of exposure and disease onset, insufficient latency, or having exposure measurements that are not reliable measures of exposure during the etiologic window(s). • Direct evidence that bias was likely because the exposure was assessed using methods with poor validity. • Evidence of differential exposure misclassification (e.g., differential recall of self-reported exposure). • There is evidence that an insufficient number of the exposure data measurements were above the limit of quantification for the assay.

[a]Reverse causality refers to a situation where an observed association between exposure and outcome is not due to causality from exposure to outcome, but rather due to the outcome of interest causing a change in the measured exposure.

In addition, there are PFAS-specific considerations for the evaluation of confounding. As discussed in Section 2.4.3, confounding across PFAS is an important area of uncertainty when interpreting the results of epidemiology studies for individual PFAS (i.e., quantifying the effected of an individual PFAS can potentially be confounded by other PFAS). Based on preliminary analyses, correlations differ across the PFAS (see Figure 6). While some pairs have correlation coefficients consistently above 0.6 (e.g., PFNA and PFDA), the correlations for most vary from 0.1 to 0.6 depending on the study, and little data is available on correlations with less commonly occurring or detected PFAS like PFBA and PFHxA.

	PFBA	PFDA	PFHxA	PFHxS	PFNA	PFOA	PFOS
PFBA*	1.00	0.01	0.45	0.14	0.15	0.03	0.07
PFDA		1.00	−0.03	0.28	0.73	0.42	0.48
PFHxA*			1.00	0.08	−0.07	0.19	−0.04
PFHxS				1.00	0.35	0.43	0.50
PFNA					1.00	0.54	0.51

Figure 6. Preliminary mean correlation coefficients across per- and polyfluoroalkyl substances (PFAS) among studies in the inventory, for all media types.

*PFBA and PFHxA correlations were based on three studies for PFOS, PFOA, and PFHxS, two studies for each other, and one study for PFNA and PFDA, so these estimates are less stable than the other PFAS, which were all based on >10 studies.
PFOS = perfluorooctane sulfonate.

Rather than rating each study with lower confidence because of this issue, potential confounding by other PFAS will be considered during the evidence synthesis phase. This may include looking across studies in populations with different exposure profiles (e.g., observing an association in a population with much higher exposure to one PFAS due to proximity to an industrial plant would increase confidence for that PFAS). In situations where there is considerable uncertainty regarding the impact of residual confounding across PFAS, this will be captured as a factor that decreases evidence strength (see Section 10).

6.3. EXPERIMENTAL ANIMAL STUDY EVALUATION

Using the principles described in Section 6.1, the evaluation of animal studies of health effects to assess risk of bias and sensitivity will be conducted for the following domains: reporting quality, risk of bias (selection or performance bias, confounding/variable control, and reporting or attrition bias), study sensitivity (exposure methods sensitivity, and outcome measures and results display) (see Table 13). Several additional considerations specific to assessing these five PFAS are outlined in Section 6.3.1.

The rationale for judgments will be documented clearly and consistently at the outcome level. In addition, similar to the evaluation of epidemiology studies, for domains other than reporting quality the evaluation documentation in HAWC will include the identified limitations and consider their impact on the overall confidence level. This will, to the extent possible, reflect an interpretation of the potential influence on the outcome-specific results (including the direction and/or magnitude of influence).

Table 13. Considerations to evaluate domains from animal toxicology studies

Evaluation concern	Domain—core question	Prompting questions	General considerations
Reporting quality	**Reporting quality** Does the study report information for evaluating the design and conduct of the study for the endpoints/outcomes of interest? *Note:* *This domain is limited to reporting. Other aspects of the exposure methods, experimental design, and endpoint evaluation methods are evaluated using the domains related to risk of bias and study sensitivity.*	Does the study report the following? **Critical information** necessary to perform study evaluation: • Species, test article name, levels and duration of exposure, route (e.g., oral; inhalation), qualitative or quantitative results for at least one endpoint of interest **Important information** for evaluating the study methods: • Test animal: strain, sex, source, and general husbandry procedures • Exposure methods: source, purity, method of administration • Experimental design: frequency of exposure, animal age and lifestage during exposure and at endpoint/outcome evaluation • Endpoint evaluation methods: assays or procedures used to measure the endpoints/outcomes of interest	A judgment and rationale for this domain will generally be given for the study. In the rationale, reviewers will also indicate when a study adhered to GLP, or to OECD (or similar) testing guidelines. • **Good:** All **critical and important information** is reported or inferable for the endpoints/outcomes of interest. • **Adequate:** All **critical information** is reported, but some **important information** is missing. However, the missing information is not expected to significantly impact the study evaluation. • **Deficient:** All **critical information** is reported, but **important information** is missing that is expected to significantly reduce the ability to evaluate the study. • **Critically deficient:** Study report is missing any pieces of **critical information**. Studies that are *Critically Deficient* for reporting are *Uninformative* for the overall rating and not considered further.

Evaluation concern		Domain—core question	Prompting questions	General considerations
Risk of bias	**Selection and performance bias**	**Allocation** Were animals assigned to experimental groups using a method that minimizes selection bias?	For each study: • Did each animal or litter have an equal chance of being assigned to any experimental group (i.e., random allocation[a])? • Is the allocation method described? • Aside from randomization, were any steps taken to balance variables across experimental groups during allocation?	A judgment and rationale for this domain will be given for each cohort or experiment in the study. • **Good**: Experimental groups were randomized, and any specific randomization procedure was described or inferable (e.g., computer-generated scheme). (Note that normalization is not the same as randomization [see response for '*Adequate*'].) • **Adequate**: Authors report that groups were randomized but do not describe the specific procedure used (e.g., "animals were randomized"). Alternatively, authors used a nonrandom method to control for important modifying factors (i.e., with respect to the outcome of interest) across experimental groups (e.g., body-weight normalization). • **Not reported** (interpreted as *Deficient*): No indication of randomization of groups or other methods (e.g., normalization) to control for important modifying factors across experimental groups. • **Critically deficient:** Bias in the animal allocations was reported or inferable.

Evaluation concern		Domain—core question	Prompting questions	General considerations
Risk of bias (continued)	**Selection and performance bias (continued)**	**Observational bias/blinding** Did the study implement measures to reduce observational bias?	For each endpoint/outcome or grouping of outcomes in a study: • Does the study report blinding or other methods/procedures for reducing observational bias, as appropriate for the assays of interest? • If not, did the study use a design or approach for which such procedures can be inferred? • What is the expected impact of failure to implement (or report implementation) of these methods/procedures on results?	A judgment and rationale for this domain will be given for each endpoint/outcome or group of outcomes investigated in the study. • **Good**: Measures to reduce observational bias were described (e.g., blinding to conceal treatment groups during endpoint evaluation; consensus-based evaluations of histopathology lesions[a]). • **Adequate**: Methods for reducing observational bias (e.g., blinding) can be inferred or were reported but described incompletely. • **Not reported:** Measures to reduce observational bias were not described. o (Interpreted as *Adequate*): The potential concern for bias was mitigated based on use of automated/computer driven systems, standard laboratory kits, relatively simple, objective measures (e.g., body or tissue weight), or screening-level evaluations of histopathology. o (Interpreted as *Deficient*): The potential impact on the results is large (e.g., outcome measures are highly subjective). • **Critically deficient:** Strong evidence for observational bias that impacted the results.

Evaluation concern		Domain—core question	Prompting questions	General considerations
Risk of bias (continued)	**Confounding/variable control**	**Confounding** Are variables with the potential to confound or modify results controlled for and consistent across all experimental groups?	For each study: • Are there differences across the treatment groups (e.g., co-exposures, vehicle, diet, palatability, husbandry, health status, surgery) that could bias the results? • If differences are identified, to what extent are they expected to impact the results?	A judgment and rationale for this domain will be given for each cohort or experiment in the study, noting when the potential for confounding is restricted to specific endpoints/outcomes. • **Good**: Outside of the exposure of interest, variables that are likely to confound or modify results appear to be controlled for and consistent across experimental groups. • **Adequate**: Some concern that variables likely to confound or modify the results were uncontrolled or inconsistent across groups, but these are expected to have a minimal impact on the results. • **Deficient**: Notable concern that potentially confounding variables were uncontrolled or inconsistent across groups and that they are expected to substantially impact the results. • **Critically deficient**: Confounding variables were presumed to be uncontrolled or inconsistent across groups, and they are expected to be a primary driver of the results.

Evaluation concern		Domain—core question	Prompting questions	General considerations
Risk of bias (continued)	**Selective reporting and attrition bias**	**Selective reporting and attrition** Did the study report results for all prespecified outcomes and tested animals? *Note:* *This domain does not consider the appropriateness of the comparisons/results presentation. This aspect of study quality is evaluated in another domain.*	For each study: Selective reporting bias: • Are all results presented for endpoints/outcomes described in the methods (see note)? Attrition bias: • Do the results account for all animals? • If there are discrepancies, do the authors provide an explanation (e.g., death or unscheduled sacrifice during the study)? • If unexplained results omissions and/or attrition are identified, what is the expected impact on the interpretation of the results?	A judgment and rationale for this domain will be given for each cohort or experiment in the study. • **Good:** Quantitative or qualitative results were reported for all prespecified outcomes (explicitly stated or inferred), exposure groups and evaluation time points. Data not reported in the primary article are available from supplemental material. If results omissions or animal attrition are identified, the authors provide an appropriate explanation, and the omissions or attrition are not expected to impact the interpretation of the results. • **Adequate:** Quantitative or qualitative results are reported for most prespecified outcomes (explicitly stated or inferred), exposure groups, and evaluation time points. Omissions and/or attrition are not explained, but they are not expected to significantly impact the interpretation of the results. • **Deficient:** Quantitative or qualitative results are missing for many prespecified outcomes (explicitly stated or inferred), exposure groups and evaluation time points, or there is high animal attrition; omissions and/or attrition are not explained and are expected to significantly impact the interpretation of the results. • **Critically deficient:** Extensive results omission and/or animal attrition are identified and prevent comparisons of results across treatment groups.

Evaluation concern		Domain—core question	Prompting questions	General considerations
Sensitivity	**Exposure methods sensitivity**	**Chemical administration and characterization** Did the study adequately characterize exposure to the chemical of interest and the exposure administration methods? *Note:* *These considerations are limited to oral exposure, as only a single inhalation study focusing on acute toxicity (i.e., after PFNA exposure) was identified (see Section 2.3.2).*	For each study: • Does the study report the source and purity and/or composition (e.g., identity and percent distribution of different isomers) of the chemical? If not, can the purity and/or composition be obtained from the supplier (e.g., as reported on the website)? • Was independent analytical verification of the test article purity and composition performed? • Are there concerns about the methods used to administer the chemical (e.g., gavage volume)? • If necessary, based on consideration of chemical-specific knowledge (e.g., instability in solution; volatility) and/or exposure design (e.g., the frequency and duration of exposure), were the chemical concentrations in the dosing solutions or diet analytically confirmed?	A judgment and rationale for this domain will be given for each cohort or experiment in the study. • **Good:** Chemical administration and characterization is complete (i.e., source, purity, and analytical verification of the test article are provided). There are no concerns about the composition, stability, or purity of the administered chemical, or the specific methods of administration. • **Adequate:** Some uncertainties in the chemical administration and characterization are identified, but these are expected to have minimal impact on interpreting the results (e.g., source and vendor-reported purity are presented, but not independently verified; purity of the test article is suboptimal but not concerning). • **Deficient:** Uncertainties in the exposure characterization are identified and expected to substantially impact the results (e.g., source of the test article is not reported; levels of impurities are substantial or concerning; deficient administration methods, such as use of a gavage volume considered too large for the species and/or lifestage at exposure). • **Critically deficient:** Uncertainties in the exposure characterization are identified, and there is reasonable certainty that the results are largely attributable to factors other than exposure to the chemical of interest (e.g., identified impurities are expected to be a primary driver of the results).

Evaluation concern		Domain—core question	Prompting questions	General considerations
Sensitivity (continued)	**Exposure methods sensitivity (continued)**	**Exposure timing, frequency, and duration** Was the timing, frequency, and duration of exposure sensitive for the endpoint(s)/outcome(s) of interest?	For each endpoint/outcome or grouping of outcomes in a study: • Does the exposure period include the full critical window of sensitivity, based on current biological understanding? • Was the duration and frequency of exposure sensitive for detecting the endpoint of interest?	A judgment and rationale for this domain will be given for each endpoint/outcome or group of outcomes investigated in the study. • **Good:** The duration and frequency of the exposure was sensitive, and the exposure included the critical window of sensitivity (if known). • **Adequate:** The duration and frequency of the exposure was sensitive, and the exposure covered most of the critical window of sensitivity (if known). • **Deficient:** The duration and/or frequency of the exposure is not sensitive and did not include most of the critical window of sensitivity (if known). These limitations are expected to bias the results towards the null. • **Critically deficient:** The exposure design was not sensitive and is expected to strongly bias the results towards the null. The rationale should indicate the specific concern(s).

Evaluation concern		Domain—core question	Prompting questions	General considerations
Sensitivity (continued)	**Outcome measures and results display**	**Endpoint sensitivity and specificity** Are the procedures sensitive and specific for evaluating the endpoint(s)/outcome(s) of interest? *Note:* • *Sample size alone is not a reason to conclude an individual study is critically deficient.* • *Considerations related to adjustments/ corrections to endpoint measurements (e.g., organ weight corrected for body weight) are addressed under results presentation.*	For each endpoint/outcome or grouping of outcomes in a study: • Are there concerns regarding the sensitivity, specificity, and/or validity of the outcome measurement protocols? • Are there serious concerns regarding the sample size? • Are there concerns regarding the timing of the endpoint assessment?	A judgment and rationale for this domain will be given for each endpoint/outcome or group of outcomes investigated in the study. Examples of potential concerns include: • Selection of protocols that are insensitive or nonspecific for the endpoint of interest. • Evaluations did not include all treatment groups (e.g., only control and high dose). • Use of unreliable or invalid methods to assess the outcome. • Assessment of endpoints at inappropriate or insensitive ages, or without addressing known endpoint variation (e.g., due to circadian rhythms, estrous cyclicity). • Decreased specificity or sensitivity of the response due to the timing of endpoint evaluation, as compared with exposure (e.g., immediate endpoint assessment after exposure to chemicals with short-acting depressant or irritant effects; insensitivity due to prolonged period of non-exposure before testing).

Evaluation concern		Domain—core question	Prompting questions	General considerations
Sensitivity (continued)	**Outcome measures and results display (continued)**	**Results presentation** Are the results presented in a way that makes the data usable and transparent? *Note:* *Potential issues associated with statistical analyses will be flagged for review by EPA statisticians and possible reanalysis (if information is available to do so, any reanalysis will be transparently presented). Any remaining limitations will be discussed during evidence synthesis or dose-response analyses (depending on the identified issue).*	For each endpoint/outcome or grouping of outcomes in a study: • Does the level of detail allow for an informed interpretation of the results? • Are the data analyzed, compared, or presented in a way that is inappropriate or misleading?	A judgment and rationale for this domain will be given for each endpoint/outcome or group of outcomes investigated in the study. Examples of potential concerns include: • Non-preferred presentation (e.g., developmental toxicity data averaged across pups in a treatment group, when litter responses are more appropriate; presentation of absolute organ-weight data when relative weights are more appropriate). • Failing to present quantitative results either in tables or figures. • Pooling data when responses are known or expected to differ substantially (e.g., across sexes or lifestages). • Failing to report on or address overt toxicity when exposure levels are known or expected to be highly toxic. • Lack of full presentation of the data (e.g., presentation of mean without variance data; concurrent control data are not presented).
Overall confidence		**Overall confidence** Considering the identified strengths and limitations, what is the overall confidence rating for the endpoint(s)/outcome(s) of interest? *Note:* *Reviewers will mark studies that are rated lower than high confidence due only to low sensitivity (i.e., bias towards the null) for additional consideration during evidence synthesis. If the study is otherwise well conducted and an effect is observed, the confidence may be increased.*	For each endpoint/outcome or grouping of outcomes in a study: • Were concerns (i.e., limitations or uncertainties) related to the reporting quality, risk of bias, or sensitivity identified? • If yes, what is their expected impact on the overall interpretation of the reliability and validity of the study results, including (when possible) interpretations of impacts on the magnitude or direction of the reported effects?	The overall confidence rating considers the likely impact of the noted concerns (i.e., limitations or uncertainties) in reporting, bias and sensitivity on the results. A confidence rating and rationale will be given for each endpoint/outcome or group of outcomes investigated in the study. Confidence rating definitions are described above (see Section 6.1).

Evaluation concern	Domain—core question	Prompting questions	General considerations

[a]Several studies have characterized the relevance of randomization, allocation concealment, and blind outcome assessment in experimental studies (Hirst et al., 2014; Krauth et al., 2013; Macleod, 2013; Higgins and Green, 2011).
GLP = good laboratory practice; OECD = Organisation for Economic Co-operation and Development.

6.3.1. Animal Toxicology Study Evaluation Considerations Specific to These Five Per- and Polyfluoroalkyl Substances (PFAS)

One of the key uncertainties in these assessments has to do with the toxicokinetics of these PFAS. The apparent differences in toxicokinetics across animal species will not be addressed at the individual study level but will be considered during evidence integration (see Section 10) and is expected to be most influential when developing toxicity values for potential human health hazards (see Section 11). However, based on Table 8 (see Section 2.4.1), the clearance of some of these PFAS, and the sex-specific differences in serum half-lives, represent important considerations for potential sources of insensitivity during study evaluation. Specifically, studies that fail to account for the short serum half-lives of PFBA in female rats and mice (half-lives of ~1–3 hours; half-lives in males are close to half a day and of less concern) and PFHxA in rats and mice of both sexes (half-lives of ~0.5–1.5 hours) by including, for example, multiple daily exposures may be judged as insensitive. Half-lives in rodents for PFDA, PFNA, and PFHxS are on the order of days or longer, and so insensitivity due to short half-life in rodents does not represent a concern (note: no nonhuman primate health effect studies were identified). Similarly, given the profound apparent differences in clearance between male and female rats for PFNA (i.e., females appear to clear PFNA 25× faster), studies that examined both sexes are preferred, and any study that tested female rats only may be judged as insensitive. This consideration may also be applied, but to a lesser extent, for studies of PFHxS in rats and PFBA in mice or rats (females appear to clear these PFAS ~4–6× faster than males).

These five PFAS are considered stable and nonreactive, and the presence of potentially toxic impurities within these readily available chemicals has not been identified as an issue in the literature. Thus, failure to describe preparation and storage of dosing solutions will not be considered an issue of concern, and a lack of information on chemical purity will not be considered a significant limitation. However, given the more relevant possibility of contamination of these five PFAS with other PFAS, a lack of analytical verification of the test article will be flagged as a limitation, although this alone will not significantly affect overall study confidence ratings.

A wide variety of outcomes have been assessed in the available animal studies for these five PFAS. Considerations specific to each outcome are not included in this protocol (outcome-specific concerns will be available in HAWC when the assessments are released). As examples, a few specific considerations that will be applied include better domain ratings for studies that address potential differences in time of day) for evaluations of hormone levels (due to fluctuations

throughout the day) and for studies that address fasting status for metabolic-related measurements.

6.4. PHARMACOKINETIC MODEL EVALUATION

A similar approach for evaluation will be applied to the full PBPK models for PFHxS (Kim et al., 2018), and for PFDA and PFNA (Kim et al., 2019), as well as to the two-compartment PK model for gestational and lactational transfer of PFHxS in humans described by Verner et al. (2016). Models will be preferred for use in these assessments when an applicable one exists and no equal or better alternative for dosimetric extrapolation is available. Given these preferences, sound justification will be provided for *not* using a PBPK (or classical PK) model when an applicable one exists and no equal or better alternative for dosimetric extrapolation is available. Note, however, that these preferences *only* apply to models that faithfully represent current scientific knowledge and accurately translate the science into computational code in a reproducible, transparent manner. In practice, it has been found that many models have errors that affect their predictions to varying degrees; hence, an evaluation of a model is required before it can be used in an assessment. Thus, the models identified above and any other models identified at later stages of developing these assessments will be evaluated as described below.

Considerations for judging the suitability of a model are separated into two categories: scientific and technical. In summary, the scientific criteria focus on whether the biology, chemistry, and other information available for chemical mode(s) of action (MOA[s]) are appropriately represented by the model structure and equations. Significant to the overall efficiency of this process, the scientific criteria can be judged by reading the publication or report that describes the model, without requiring an evaluation of the computer code. Preliminary technical criteria include the availability of the computer code and apparent completeness of parameter listing and documentation. The in-depth technical and scientific criteria focus on the accurate implementation of the conceptual model in the computational code, use of correct or biologically consistent parameters in the model, and reproducibility of model results reported in journal publications and other documents. Specific details for this evaluation are provided in the Quality Assurance Project Plan for PBPK models (U.S. EPA, 2018b).

6.5. MECHANISTIC STUDY EVALUATION

Sections 9 and 10 outline an approach for focused consideration of information from mechanistic studies (including *in vitro*, *in vivo*, *ex vivo*, and *in silico* studies) where the specific analytical approach is targeted to the assessment needs, depending in part on the extent and nature of the phenotypic human and animal evidence. In this way, the mechanistic synthesis for a given health effect might range from a high-level summary (or detailed analysis) of potential mechanisms of action to specific, focused questions needed to address important and impactful assessment uncertainties unaddressed by the available phenotypic studies (e.g., expected shape of the

dose-response curve in the low-dose region, applicability of the animal evidence to humans, addressing susceptible populations). Individual study-level evaluation of mechanistic endpoints will not typically be pursued. However, it may be necessary to identify assay-specific considerations for study endpoint evaluations on a case-by-case basis to provide a more detailed summary and evaluation for the most relevant individual mechanistic studies addressing a key assessment uncertainty. This may be done, for example, when the scientific understanding of a critical mechanistic event or MOA lacks scientific consensus, when the reported findings on a critical mechanistic endpoint are conflicting, when the available mechanistic evidence addresses a complex and influential aspect of the assessment, or when *in vitro* or *in silico* data make up the bulk of the evidence base and there is little or no evidence from epidemiological studies or animal bioassays. As noted in Section 3 and Section 4, genotoxicity studies were identified as meeting PECO criteria; these data will be summarized in each PFAS assessment to describe evidence relevant to carcinogenicity even in the absence of more phenotypic data. Based on the considerations above, if these studies are interpreted as adequate to draw a hazard conclusion (i.e., other than "insufficient"), individual study-level evaluations of some or all the genotoxicity studies will be informative to this decision. If necessary, based on the assessment-specific issues identified during study evaluation and evidence synthesis (see Section 9.2), the specific approach to evaluating individual studies other than those addressed in Sections 6.2–6.4 will be outlined in the assessments and included as an update to the protocol.

7. ORGANIZING THE HAZARD REVIEW

The organization and scope of the hazard evaluation is determined by the available evidence for each PFAS regarding routes of exposure, metabolism and distribution, outcomes evaluated, and number of studies pertaining to each outcome, as well as the results of the evaluation of sources of bias and sensitivity. The hazard evaluations will be organized around organ systems (e.g., nervous system) informed by one or multiple related outcomes, as described in Section 5, and a decision will be made as to what level (e.g., organ system or subsets of outcomes within an organ system) to organize the synthesis.

Table 14 lists some questions that may be asked of the evidence to assist with this decision. These questions extend from considerations and decisions made during development of the refined evaluation plan to include review of the concerns raised during individual study evaluations as well as the direction and magnitude of the study-specific results. Resolution of these questions will then inform critical decisions about the organization of the hazard evaluation and help determine what studies may be useful in dose-response analyses.

Table 14. Querying the evidence to organize syntheses for human and animal evidence

Evidence	Questions	Follow-up questions
ADME	Given the known ADME issues for these PFAS, do the data appear to differ by route of exposure studied, lifestage when exposure occurred, sex, species, or dosing regimens used?	Will separate analyses be needed by factors such as sex, route of exposure, or by methods of dosing within a route of exposure (e.g., are large differences expected between gavage and dietary exposures)? Are data available to inform which lifestages and what dosing regimens are more relevant to human exposure scenarios?
	Is there toxicity information for metabolites that also should be evaluated for hazard?	What exposures will be included in the evaluation?
Outcomes	What outcomes are reported in studies? Are the data reported in a comparable manner across studies (similar output metrics at similar levels of specificity, such as adenomas and carcinomas quantified separately)?	At what level (hazard, grouped outcomes, or individual outcomes) will the synthesis be conducted? What commonalities will the outcomes be grouped by: • health effect, • exposure levels, • functional or population-level consequences (e.g., endpoints all ultimately leading to decreased fertility or impaired cognitive function), • involvement of related biological pathways How well do the assessed human and animal outcomes relate within a level of grouping?
	Are there inter-related outcomes? If so, consider whether some outcomes are more useful and/or of greater concern than others.	
	Does the evidence indicate greater sensitivity to effects (at lower exposure levels or severity) in certain subgroups (by age, sex, ethnicity, lifestage)? Should the hazard evaluation include a subgroup analysis?	
	Does incidence or severity of an outcome increase with duration of exposure or a particular window of exposure. What exposure time frames are relevant to development or progression of the outcome?	
	Is there mechanistic evidence that informs how outcomes might be grouped together?	
	How robust is the evidence for specific outcomes? • What outcomes are reported by both human and animal studies and by one or the other? Were different animal species and sexes (or other important population-level differences) tested? • In general, what are the study confidence conclusions of the studies (*high*, *medium*, *low*, *uninformative*) for the different outcomes? Is there enough evidence from *high*- and *medium*-confidence studies to draw conclusions about causality?	What outcomes should be highlighted? Should the others be synthesized at all? Would comparisons by specific limitations be informative?

Evidence	Questions	Follow-up questions
Dose-response	Did some outcomes include better coverage of exposure ranges that may be most relevant to human exposure than others?	What outcomes and studies are informative for developing toxicity values?
	For which outcomes are there sufficient data available to draw conclusions about dose-response? Are there outcomes with study results of sufficient similarity (e.g., an established linkage in a biological pathway) to allow examination or calculation of common measures of effect across studies? Do the mechanistic data identify surrogate or precursor outcomes that are adequate for dose-response analysis?	
	Are there subgroups that exhibit responses at lower exposure levels than others?	
	Are there findings from ADME studies that could inform data-derived extrapolation factors, or link toxicity observed via different routes of exposure, or link effects between humans and experimental animals?	What studies might be used to develop non-default UFs? Is there a common internal dose metric that can be used to compare species or routes of exposure?

ADME = absorption, distribution, metabolism, and excretion; UF = uncertainty factor.

1

8. DATA EXTRACTION OF STUDY METHODS AND RESULTS

Data extraction and content management will be carried out using the Health Assessment Workplace Collaborative (HAWC; web links will be shared in the individual assessments). A consistent approach to data extraction will be applied across these PFAS assessments to facilitate their anticipated future use in addressing poorly studied PFAS (e.g., through coupling with computational toxicology data generated as described in the Environmental Protection Agency [EPA] PFAS action plan). Data extraction elements that may be collected from epidemiological, controlled human exposure, animal toxicological, and *in vitro* studies are described in HAWC (https://hawcprd.epa.gov/about/). Not all studies that meet the PECO criteria go through data extraction. Studies evaluated as being *Uninformative* are not considered further and therefore would not undergo data extraction. In addition, outcomes determined to be less relevant during PECO refinement (see Section 5) may not go through data extraction or may have only minimal data extraction. The same may be true for *low*-confidence studies if enough *medium*- and *high*-confidence studies (e.g., on an outcome) are available. All findings are considered for extraction, regardless of statistical significance. The level of extraction for specific outcomes within a study may differ (i.e., ranging from a narrative to full extraction of dose-response effect size information). In part, this is determined based on the level of detail to be discussed in the evidence synthesis for that health effect (e.g., a detailed extraction will not be necessary for health effects with very few available studies; these will only be briefly summarized in a short narrative). Similarly, decisions about data extraction for *low*-confidence studies are typically made while implementing the protocol and are based on the quality and extent of the available evidence. If necessary, the version of the protocol released with the draft assessment will outline how *low*-confidence studies were treated for extraction and evidence synthesis.

The data extraction results for included studies will be presented in the assessment (and made available for download from EPA HAWC in Excel format) when the draft is publicly released. (Note: The following browsers are supported for accessing HAWC: Google Chrome [preferred], Mozilla Firefox, and Apple Safari. There are errors in functionality when viewed with Internet Explorer.) For quality control, data extraction will be performed by one member of the evaluation team and independently verified by at least one other member. Discrepancies in data extraction will be resolved by discussion or consultation with a third member of the evaluation team. Digital rulers, such as WebPlotDigitizer (https://automeris.io/WebPlotDigitizer/), will be used to extract numerical information from figures, and their use will be documented during extraction.

As previously described, routine attempts will be made to obtain missing information from human and animal health effect studies, if it is considered influential during study evaluations (see Section 6) or when it can provide information important for dose-response analysis or interpretations of significance (e.g., missing group size or variance descriptors such as standard deviation or confidence interval). Missing data from individual mechanistic (e.g., *in vitro*) studies generally will not be sought. Outreach to study authors or designated contact persons will be documented and considered unsuccessful if researchers do not respond to email or phone requests within 1 month of initial attempt(s) to contact.

8.1. STANDARDIZING REPORTING OF EFFECT SIZES

In addition to providing quantitative outcomes in their original units for all study groups, results from outcome measures will be transformed, when possible, to a common metric to help compare distinct but related outcomes that are measured with different scales. These standardized effect size estimates facilitate systematic evaluation and evidence integration for hazard identification (see Section 9.1). The following summary of effect size metrics by data type outlines issues in selecting the most appropriate common metric for a collection of related endpoints (Vesterinen et al., 2014).

Common metrics for continuous outcomes include:

- *Absolute difference in means.* This metric is the difference between the means in the control and treatment groups, expressed in the units in which the outcome is measured. When the outcome measure and its scale are the same across all studies, this approach is the simplest to implement.

- *Percent control response (or normalized mean difference [NMD]).* Percent control group calculations are based on means. Standard deviation (or standard error) values presented in the studies for these normalized effect sizes can also be estimated if sufficient information has been provided. Note that some outcomes reported as percentages, such as mean percentage of affected offspring per litter, can lead to distorted effect sizes when further characterized as percentage change from control. Such measures are better expressed as absolute difference in means, or rather preferably transformed to incidences using approaches for event or incidence data (see below).

- *Standardized mean difference.* The NMD approach above is relevant to ratio scales, but sometimes it is not possible to infer what a "normal" animal would score, such as when data for animals without lesions are not available. In these circumstances, standardized mean differences can be used. The difference in group means is divided by a measure of the pooled variance to convert all outcome measures to a standardized scale with units of standard deviations. This approach can also be applied to data for which different measurement scales are reported for the same outcome measure (e.g., different measures of lesion size such as infarct volume and infarct area).

Common metrics for event or incidence data include:

- Percent change from control. This metric is analogous to the NMD approach described for continuous data above.

- For binary outcomes such as the number of individuals that developed a disease or died, and with only one treatment evaluated, data can be represented in a 2 × 2 table. Note that when the value in any cell is 0, 0.5 is added to each cell to avoid problems with the computation of the standard error. For each comparison, the odds ratio (OR) and its standard error can be calculated. Odds ratios are normally combined on a logarithmic scale.

An additional approach for epidemiology studies is to extract adjusted statistical estimates when possible rather than unadjusted or raw estimates.

It is important to consider the variability associated with effect size estimates, with better powered studies generally showing more precise estimates. Effect size estimation can be affected, however, by such factors as variances that differ substantially across treatment groups, or by lack of information to characterize variance, especially for animal studies in biomedical research (Vesterinen et al., 2014). The assessments will consider the nature of any variance issues and ensure that the associated uncertainties are clarified and accounted for during the evidence synthesis process (see Section 9).

8.2. STANDARDIZING ADMINISTERED DOSE LEVELS/CONCENTRATIONS

Exposures will be standardized to common units. Exposure levels in oral studies will be expressed in units of mg PFAS/kg-day. Where study authors provide exposure levels in concentrations in the diet or drinking water, dose conversions will be made using study-specific food or water consumption rates and body weights when available. Otherwise, EPA defaults will be used (U.S. EPA, 1988), addressing age and study duration as relevant for the species/strain and sex of the animal of interest. Exposure levels in inhalation studies will be expressed in units of mg/m^3. Assumptions used in performing dose conversions will be documented in HAWC or the specific assessments.

9. SYNTHESIS OF EVIDENCE

For the purposes of this assessment, evidence synthesis and integration are considered distinct, but related processes. The syntheses of separate bodies of evidence (i.e., human, animal, and mechanistic evidence) described in this section will directly inform the integration across all evidence to draw an overall judgment for each of the assessed human health effects (as described in Section 10). The phrase "evidence integration" used here is analogous to the phrase "weight of evidence" used in some other assessment processes (EFSA, 2017; U.S. EPA, 2017a; NRC, 2014; U.S. EPA, 2005a).[12]

For each potential human health effect (or smaller subset of related outcomes), the U.S. Environmental Protection Agency (EPA) will separately synthesize the available phenotypic human and animal evidence pertaining to that potential health effect. Mechanistic evidence will also be considered in targeted analyses conducted prior to, during, and after developing syntheses of the phenotypic human and animal evidence. The results of the analyses of mechanistic evidence will be used to inform key uncertainties; as a result, the scope of the mechanistic analyses will generally depend on the extent and nature of the human and animal evidence (see Sections 9.2 and 10). Thus, apart from the pre-defined mechanistic analyses (see Sections 9.2.1-9.2.3), the human and animal evidence syntheses (or the lack of phenotypic data in humans and animals) help determine the approach to be taken in synthesizing the available mechanistic evidence (see Section 9.2.4). In this way, a mechanistic evidence synthesis might range from a high-level summary of potential toxicity mechanisms discussed in the published literature to a detailed analysis of multiple potential modes-of-action, or it might evaluate specific, focused questions that inform key uncertainties unaddressed by the phenotypic human and animal evidence (e.g., shape of the dose-response curve at low doses, applicability of the animal evidence to humans, addressing susceptible populations). Each synthesis will provide a summary discussion of the available evidence that addresses considerations regarding causation (see Table 15). These considerations are adapted from considerations for causality introduced by Austin Bradford Hill (Hill, 1965): consistency, dose-response relationship, strength of the association, temporal relationship, biological plausibility, coherence, and "natural experiments" in humans [see additional discussion

[12] This revision has been adopted primarily based on the 2014 NAS review of IRIS (NRC, 2014): "The present committee found that the phrase *weight of evidence* has become far too vague as used in practice today and thus is of little scientific use. In some accounts, it is characterized as an oversimplified balance scale on which evidence supporting hazard is placed on one side and evidence refuting hazard on the other... The present committee found the phrase *evidence integration* to be more useful and more descriptive of what is done at this point in an IRIS assessment—that is, IRIS assessments must come to a judgment about whether a chemical is hazardous to human health and must do so by integrating a variety of evidence."

in U.S. EPA (2005a) and U.S. EPA (1994)]. Importantly, the evidence synthesis process explicitly considers and incorporates the conclusions from the individual study evaluations (see Section 6).

Table 15. Information most relevant to describing primary considerations for assessing causality during evidence syntheses

Consideration	Description of the consideration and its application in IRIS syntheses
Study confidence	Description: Incorporates decisions about study confidence within each of the considerations. Application: In evaluating the evidence for each of the causality considerations described in the following rows, the syntheses will consider study confidence decisions. *High*-confidence studies carry the most weight. The syntheses will consider the specific limitations and strengths identified during study evaluation and describe how these informed each consideration.
Consistency	Description: Examines the similarity of results (e.g., direction; magnitude) across studies. Application: Syntheses will evaluate the homogeneity of findings on a given outcome or endpoint across studies. When inconsistencies exist, the syntheses consider whether results were "conflicting" (i.e., unexplained positive and negative results in similarly exposed human populations or in similar animal models) or "differing" (i.e., mixed results explainable by, for example, differences between human populations, animal models, exposure conditions, or study methods) (U.S. EPA, 2005a). These considerations are based on analyses of potentially important explanatory factors such as: • Confidence in studies' results, including study sensitivity (e.g., some study results that appear to be inconsistent may be explained by potential biases or other attributes that affect sensitivity). • Exposure, including route (if applicable) and administration methods, levels, duration, timing with respect to outcome development (e.g., critical windows), and exposure assessment methods (i.e., in epidemiology studies), including analytical units and specific groups being compared. • Specificity and sensitivity of the endpoint for evaluating the health effect in question (e.g., functional measures can be more sensitive than organ weights). • Populations or species, including consideration of potential susceptible groups or differences across lifestage at exposure or endpoint assessment. • Toxicokinetic information explaining observed differences in responses across route of exposure, other aspects of exposure, species, sexes, or lifestages. The interpretation of consistency will emphasize biological significance, to the extent that it is understood, over statistical significance. Statistical significance from suitably applied tests (this may involve consultation with an EPA statistician) adds weight when biological significance is not well understood. Consistency in the direction of results increases confidence in that association even in the absence of statistical significance. In some cases, it may be helpful to consider the potential for publication bias to provide context to interpretations of consistency.[a]

Consideration	Description of the consideration and its application in IRIS syntheses
Strength (effect magnitude) and precision	Description: Examines the effect magnitude or relative risk, based on what is known about the assessed endpoint(s), and considers the precision of the reported results based on analyses of variability (e.g., confidence intervals; standard error). This may include consideration of the rarity or severity of the outcomes. Application: Syntheses will analyze results both within and across studies and may consider the utility of combined analyses (e.g., meta-analysis). While larger effect magnitudes and precision (e.g., $p < 0.05$) help reduce concerns about chance, bias, or other factors as explanatory, syntheses should also consider the biological or population-level significance of small effect sizes.
Biological gradient/ dose-response	Description: Examines whether the results (e.g., response magnitude; incidence; severity) change in a manner consistent with changes in exposure (e.g., level; duration), including consideration of changes in response after cessation of exposure. Application: Syntheses will consider relationships both within and across studies, acknowledging that the dose-response relationship (e.g., shape) can vary depending on other aspects of the experiment, including the biology underlying the outcome and the toxicokinetics of the chemical. Thus, when dose-dependence is lacking or unclear, the synthesis will also consider the potential influence of such factors on the response pattern.
Coherence	Description: Examines the extent to which findings are cohesive across different endpoints that are related to, or dependent on, one another (e.g., based on known biology of the organ system or disease, or mechanistic understanding such as toxicokinetic/dynamic understanding of the chemical or related chemicals). In some instances, additional analyses of mechanistic evidence from research on the chemical under review or related chemicals that evaluate linkages between endpoints or organ-specific effects may be needed to interpret the evidence. These analyses may require additional literature search strategies. Application: Syntheses will consider potentially related findings, both within and across studies, particularly when relationships are observed within a cohort or within a narrowly defined category (e.g., occupation; strain or sex; lifestage of exposure). Syntheses will emphasize evidence indicative of a progression of effects, such as temporal- or dose-dependent increases in the severity of the type of endpoint observed. If an expected coherence between findings is not observed, possible explanations should be explored including those related to the biology of the effects as well as the sensitivity and specificity of the measures used.

Consideration	Description of the consideration and its application in IRIS syntheses
Mechanistic evidence related to biological plausibility	Description: There are multiple uses for mechanistic information, and this consideration overlaps with "coherence." This consideration examines the biological support (or lack thereof) for findings from the human and animal health effect studies and becomes more influential on the hazard conclusions when notable uncertainties in the strength of those sets of studies exist. These analyses can also improve understanding of dose- or duration-related development of the health effect. In the absence of human or animal evidence of apical health endpoints, the synthesis of mechanistic information may drive evidence integration conclusions (when such information is available). Application: Syntheses can evaluate evidence on precursors, biomarkers, or other molecular or cellular changes related to the health effect(s) of interest to describe the likelihood that the observed effects result from exposure. This evaluation will entail an analysis of existing evidence, and not simply whether a theoretical pathway can be postulated. This analysis may not be limited to evidence relevant to the PECO but may also include evaluations of biological pathways (e.g., for the health effect; established for other, possibly related, chemicals). Any such synthesis of mechanistic evidence will consider the sensitivity of the mechanistic changes and the potential contribution of alternative or previously unidentified mechanisms of toxicity.
Natural experiments	Description: Specific to epidemiology studies and rarely available, this consideration examines effects in populations that have experienced well-described, pronounced changes in chemical exposure (e.g., lead exposures before and after banning lead in gasoline). Application: Compared to other observational designs, natural experiments have the benefit of dividing people into exposed and unexposed groups without them influencing their own exposure status. During synthesis, associations in *medium*- and *high*-confidence natural experiments can substantially reduce concerns about residual confounding.

[a]Publication bias involves the influence of the direction, magnitude, or statistical significance of the results on the likelihood of a paper being published; it can result from decisions made, consciously or unconsciously, by study authors, journal reviewers, and journal editors (Dickersin, 1990). When evidence of publication bias is present for a set of studies, less weight may be placed on the consistency of the findings for or against an effect during evidence synthesis and integration.

PECO = populations, exposures, comparators, and outcomes.

Data permitting, the syntheses will also discuss analyses relating to potential susceptible populations.[13] These analyses will be based on knowledge about the health outcome or organ system affected, demographics, genetic variability, lifestage, health status, behaviors or practices, and social determinants (see Table 16). This information will be used to draw conclusions

[13]Various terms have been used to characterize populations that may be at increased risk of developing health effects from exposure to environmental chemicals, including "susceptible," "vulnerable," and "sensitive." Furthermore, these terms have been inconsistently defined across the scientific literature. The term susceptibility is used in this protocol to describe populations or lifestages at increased risk, focusing on intrinsic biological factors that can modify the effect of a specific exposure, but also considering social determinants or behaviors that may increase susceptibility. However, factors resulting in higher exposures to specific groups (e.g., proximity, housing, occupation) will typically not be analyzed to describe increased risk among specific populations or subgroups.

regarding potential susceptibility among specific populations or subgroups in a separate section (see Section 10.3). This summary will describe concerns across the available evidence for all potential human health effects and will inform both hazard identification and dose-response analyses.

Table 16. Individual and social factors that may increase susceptibility to exposure-related health effects

Factor	Examples
Demographic	Gender, age, race/ethnicity, education, income, occupation, geography
Genetic variability	Polymorphisms in genes regulating cell cycle, DNA repair, cell division, cell signaling, cell structure, gene expression, apoptosis, and metabolism
Lifestage	In utero, childhood, puberty, pregnancy, women of childbearing age, old age
Health status	Pre-existing conditions or disease such as psychosocial stress, elevated body mass index, frailty, nutritional status, chronic disease
Behaviors or practices	Diet, mouthing, smoking, alcohol consumption, pica, subsistence or recreational hunting and fishing
Social determinants	Income, socioeconomic status, neighborhood factors, health care access, and social, economic, and political inequality

EPA ExpoBox Exposure Assessment Tools, based on EPA's *Guidelines for Exposure Assessment* (U.S. EPA, 1992).
DNA = deoxyribonucleic acid.

9.1. SYNTHESES OF HUMAN AND ANIMAL HEALTH EFFECTS EVIDENCE

The syntheses of the human and animal health effects evidence will focus on describing aspects of the evidence that best inform causal interpretations, including the exposure context examined in the sets of studies. Each evidence synthesis will be based primarily on studies of *high* and *medium* confidence. *Low*-confidence studies may be used if few or no studies with higher confidence are available to help evaluate consistency, or if the study designs of the *low*-confidence studies address notable uncertainties in the set of *high*- or *medium*-confidence studies on a given health effect. If *low*-confidence studies are used, then a careful examination of risk bias and sensitivity with potential impacts on the evidence synthesis conclusions will be included in the narrative.

As previously described, these syntheses will articulate the strengths and the weaknesses of the available evidence organized around the considerations described in Table 15, as well as issues that stem from the evaluation of individual studies (e.g., concerns about bias or sensitivity). If possible, results across studies will be compared using graphs and charts or other data visualization strategies. The analysis will typically include examination of results stratified by any

or all of the following: study confidence classification (or specific issues within confidence evaluation domains), population or species, exposures (e.g., level, patterns [intermittent or continuous], duration, intensity), sensitivity (e.g., low vs. high), and other factors that may have been identified during study evaluation or analyses of key science issues (see Section 2.4). The number of studies and the differences encompassed by the studies will determine the extent to which specific factors can be examined for use in stratifying study results. Additionally, for both the human and animal evidence syntheses, if supported by the available data, additional analyses across studies (such as meta-analysis) may also be conducted.

9.2. MECHANISTIC INFORMATION

The synthesis of mechanistic information informs the integration of health effects evidence for both hazard identification (i.e., biological plausibility or coherence of the available human or animal evidence; inferences regarding human relevance, or the identification of susceptible populations and lifestages across the human and animal evidence) and dose-response evaluation. As introduced in prior sections, several key science issues that are essential to consider in these five assessments will involve a focused analysis and synthesis of mechanistic information (see Sections 9.2.1–9.2.3). Other potential assessment-specific uncertainties for which mechanistic analyses might be conducted, and the considerations for including those analyses in an assessment, are outlined in Section 9.2.4. Deviations from the approaches described in Sections 9.2.1–9.2.3, as well as the specific methods for any analyses conducted based on the considerations described in Section 9.2.4, will be tracked as updates to the protocol.

Mechanistic evidence includes any experimental measurement related to a health outcome that provides information about the biological or chemical events associated with phenotypic effects; these measurements can improve understanding of the mechanisms involved in the toxic effects following exposure to a chemical but are not generally considered adverse outcomes. Mechanistic data are reported in a diverse array of observational and experimental studies across species, model systems, and exposure paradigms, including *in vitro*, *in vivo* (by various routes of exposure), *ex vivo*, and *in silico* studies, and across a wide spectrum of diverse endpoints.

Evaluations of mechanistic information typically differ from evaluations of phenotypic evidence (e.g., from routine toxicology studies). This is primarily because mechanistic data evaluations consider the support for and involvement of specific events or sets of events within the context of a broader research question (e.g., support for a hypothesized mechanism; consistency with known biological processes), rather than evaluations of individual apical endpoints considered in relative isolation. Such analyses are complicated because a chemical may operate through multiple mechanistic pathways, even if one hypothesis dominates the literature (U.S. EPA, 2005a). Similarly, multiple mechanistic pathways might interact to cause an adverse effect. Thus, pragmatic and stepwise approaches to considering and reviewing this evidence for these PFAS assessments are outlined below. The format of these syntheses is expected to vary from a short narrative

summary of existing knowledge to an in-depth analysis and weighing of the evidence underlying multiple mechanistic events, depending on data availability and the criticality of the assessment-specific uncertainty(ies).

9.2.1. Toxicokinetic Information and Pharmacokinetic (PK)/Physiologically Based Pharmacokinetic (PBPK) Models

One key mechanistic issue has to do with the toxicokinetics of these chemicals, particularly their serum half-life values because these values are useful for extrapolating doses from exposed animals to humans. Toxicokinetic studies were extracted for consideration (from the broad PFAS literature searches) by subject matter experts using two different methods: (1) tagging of studies during literature screening (see Sections 4.2–4.3), noting that this tagging was not conducted by ADME subject matter experts, and (2) use of SWIFT Review software [https://www.sciome.com/swift-review/; (Howard et al., 2016)] to categorize the literature via health outcome tags for ADME from the title and abstract. For identification of ADME-related studies to be reviewed using SWIFT Active Screener (https://www.sciome.com/swift-activescreener/), the results using the health outcome tags for ADME embedded within SWIFT Review were confirmed using a search string developed by experts in toxicokinetics within the IRIS program.[14] This resulted in 813 potentially relevant studies that were imported into SWIFT Active Screener for review by two independent reviewers with demonstrated expertise in ADME (conflicts were resolved through discussion). A basic set of PECO criteria were used for this review:

- Population: *in vivo* studies in humans, nonhuman primates, rats, or mice. (Note: *in vitro* studies in these species were tagged as potentially supportive; see explanation below.)

- Exposure: any route of administration of a single chemical compound that is expected to occur for human exposure for PFBA, PFHxS, PFHxA, PFDA, or PFNA. Exposure to metabolic precursors of these chemicals was also included. (Note: intraperitoneal [i.p.] injection studies and *in vitro* studies were tagged as potentially supportive; see explanation below.)

[14]tiab: (adme OR admet OR bile OR biliary OR bioavail* OR biodistribut* OR biologic-avail* OR biological-avail* OR biologically-avail* OR biotrans* OR clearance OR detox* OR distribut* OR dosim* OR eliminat* OR endocytosis OR enterohepatic OR "entero hepatic" OR excret* OR exhalation OR hepatobiliary OR inhalation OR metaboli* OR "partition coefficient" OR permeability OR persistence OR phagocytosis OR pharmacokinetic* OR physiologic-avail* OR physiological-avail* OR physiologically-avail* OR pinocytosis OR protein-bind* OR reabsorption OR retention OR secretion* OR toxicokinetic* OR transport OR uptake OR urination OR ((absorb OR absorbs OR absorbed OR absorption* OR deposition) NOT (atomic OR optical OR spectra* OR spectros* OR spectrum* OR infrared)) OR title : ("gas exchange" AND (alveolar OR lung OR lungs OR pulmonary OR respirat*)) OR mesh_mh: ("biological transport" OR "enterohepatic circulation" OR pharmacokinetics) OR mesh_sh: (pharmacokinetics) OR mesh_mh: (toxicokinetics)).
This string identified two fewer potentially relevant studies than the SWIFT review (including all studies identified using the non-SWIFT string). So, the studies identified by SWIFT Review were screened in SWIFT Active Screener.

- Comparator: vehicle control or reference population.
- Outcome: data to quantify ADME processes, steady state analysis, empirical pharmacokinetic (PK), full PBPK.

This screening (i.e., to 96% predicted completion based on the machine-learning software) resulted in the identification of 99 studies relevant to toxicokinetics across the five PFAS assessments. These data will be considered for use in the assessments as described below.

All PK and PBPK models will be formally evaluated for use in the assessments, as described in Section 6.4. The specific approaches for determining the most appropriate method for dosimetric extrapolation, if necessary for these assessments (note: this is likely to be necessary, based on the preliminary literature inventory), as well as other potential quantitative approaches for using the PK/PBPK models and ADME data, are outlined in Section 11.2.

To draw conclusions regarding the most appropriate serum half-life measures, the ADME studies identified by the screening methods described above will be considered as outlined in U.S. EPA (2018b). Briefly, the studies relevant to updating the data presented in Table 7 in Section 2.4.1, including the studies underlying the current data in the table, will be reviewed, and data that are highly unreliable will be excluded (e.g., data points below the limit of detection [LOD]; values based on uncertain exposure estimates, or other unvalidated assumptions). Study characteristics that will be reviewed by subject matter experts to determine whether studies are informative to the PFAS-specific half-life values include appropriateness of the analytic method, the number of exposure levels tested, the human relevance of the exposure range, and the number of time points and tissues sampled. Although ADME data from *in vitro* studies and i.p. injection studies were tracked as potentially relevant during screening, additional considerations will apply to the potential incorporation of these data into the assessments, given their inherent uncertainties (e.g., difficulties interpreting the relevance of bioavailability or peak concentration data from i.p. injection studies). Specifically, regarding *in vitro* studies, it is expected that there may be no *in vivo* toxicokinetic data on the rate of conversion of precursor compounds to the PFAS of interest, in which case conversion rates measured *in vitro* can be extrapolated to *in vivo* as the next best means of predicting this mechanism of exposure. Even if such extrapolation is determined to be quantitatively uncertain, these data might still provide useful qualitative information.

In general, when there are multiple studies informative to a given ADME parameter, if the values from two or more studies of the same species, strain, and sex are similar, a numerical average among those values will be used. Because significant differences in the half-life between males and females of a given species have been observed for some PFAS, these sex differences will be assumed to be real in general across species. Specifically, when feasible, the data for males and females of each species, for each PFAS, will be analyzed separately, even if the difference is not statistically significant. If the values differ significantly across studies for the same

species/strain/sex, a more detailed review of the study methods indicated above will be conducted to determine whether one study is more likely to provide accurate information than another.

Given that PFAS inhalation exposure is expected to be via adsorption to particulates, if sufficient data are available for any of the assessed PFAS, inhalation exposure rates for PFAS-containing particles will be predicted using the multiple-path particle dosimetry (MPPD) model (https://www.ara.com/products/multiple-path-particle-dosimetry-model-mppd-v-304). This model predicts inhaled particle deposition in laboratory animals and humans as a function of particle size. Particle sizes used in controlled animal studies or measured in ambient environmental or workplace exposure studies in humans will be used as inputs. If PK data are identified that allow the bioavailability of inhaled particulate PFAS to be estimated, the mass deposition predicted by the MPPD model will be adjusted accordingly. Otherwise 100% absorption of PFAS from inhalation deposition will be assumed. Note that while some inhaled particles are later moved by the mucociliary apparatus to the larynx and swallowed, the PK bioavailability for oral ingestion can then be applied to that fraction. Any predictions will be considered for use in comparing findings across oral and inhalation routes of exposure during evidence integration (see Section 10). In addition, see Section 11.2 for the application of these predictions to developing quantitative estimates. If necessary, inhalation of PFAS in free ionic form can be estimated based on the inhalation uptake of other chemicals with high liquid: air partition coefficients (i.e., assuming nearly complete absorption of any free ions that contact the airway lining).

9.2.2. Peroxisome Proliferator-Activated Receptor Alpha (PPARα) Dependence for Health Effect(s) Observed in Animals

A second area of focused mechanistic analysis is evaluating the human relevance of effects in animals that appear to involve (at least in part) a PPARα-mediated MOA. The approach outlined below focuses on hepatic effects, which are expected to be the primary health effect area in these assessments for which this analysis is useful, and for which data are likely to exist to conduct an analysis. The specifics of applying this approach may vary across the five PFAS assessments, depending on the availability of data to address this question and the strength of the evidence indicating PPARα involvement. During assessment development, for other health effects with evidence that a PPARα-mediated MOA might be operant, the mechanistic syntheses will include consideration of this issue. These analyses will depend on the amount of information available and the strength of the evidence indicating PPARα involvement. Thus, the analyses might range from a short summary of the available evidence when data are sparse to an evaluation approximating the one described below when extensive data are available.

To identify the literature most relevant to addressing the question of the PPARα-dependence of hepatic effects observed in experimental animals, a PFAS assessment with

extensive evidence of liver effects and potential PPARα involvement will screen[15] the "potentially relevant supplemental material" studies on a given PFAS at the full-text level as follows:

- Population: *in vivo* animal studies in mammalian models; *in vitro* and human experiments using primary or immortalized liver cell lines
- Exposure: PFAS of interest (parent compound only)
- Comparator: vehicle control
- Outcome: mechanistic outcomes relevant to the hepatobiliary system (e.g., in liver tissues or cells)

Any additional assessment-specific strategies for identifying other information of potential relevance on molecular mechanistic data for these five PFAS, or from the more extensive literature on perfluorooctanoic acid (PFOA) and PFAS (e.g., as points of comparison), will be described as updates to this protocol.

The pool of studies identified based on the strategies outlined above will be inventoried into a database to allow for the organization and evaluation of these data. Specifically, the following information will be extracted for each reference: a reference identifier; test compound; exposure route and duration; the sex, species, and strain of the organism; age at exposure; and endpoint evaluation of the test organism or test system. Additionally, the inventory(ies) will capture a succinct description of the assessed endpoints and the potential mechanistic event(s) informed by those endpoints in each study. The mechanistic events in the proposed mechanisms pathway for which there are data will then be organized according to the following levels of biological organization: molecular target(s), cellular response(s), tissue/organ response(s), and organism response(s), in accordance with the levels of biological organization used to develop adverse outcome pathways (AOPs).[16]

Although refinements based on the assessment-specific evidence are anticipated, these assessments will first consider the use of the preliminary pathway outlined in Figure 7 as an organizing AOP for these data. The preliminary, proposed AOP displayed in Figure 7 is based on molecular initiating events, key events, and adverse outcomes identified in previous evaluations on PFOS and PFOA and proposed AOPs for chemical-induced noncancer liver toxicity [see Li et al.

[15]Although the specifics of this screening process may vary across PFAS, this protocol describes that screening will occur by at least two reviewers and use of DistillerSR to track decisions and resolve differences. Any deviations from this will be tracked on an assessment-specific basis as updates to the protocol.

[16]Although the World Health Organization (WHO)-International Programme on Chemical Safety (IPCS)-MOA and the Organisation for Economic Co-operation and Development (OECD)-AOP frameworks are similar in the identification and analysis of key events following modified Bradford Hill criteria (Meek et al., 2014), AOPs are chemical agnostic whereas MOA analyses are intended to inform health assessments of individual (or groups of) chemical(s) (Edwards et al., 2016).

(2017), Mellor et al. (2016), Wang et al. (2014), U.S. EPA (2016d), U.S. EPA (2016d), ATSDR (2018), and NJDWQI (2017)]. Prior evaluations of PFOS and PFOA have discussed studies using wild type, PPARα knockout and humanized PPARα (hPPARα) mice showing that exposure leads to fatty acid and triglyceride accumulation in the liver and steatosis via both PPARα-dependent and -independent pathways (ATSDR, 2018; Li et al., 2017; Viberg and Eriksson, 2017). In addition to PPARα, these reviews have implicated other nuclear receptor (NR) and cell signaling pathways with PFOA- and PFOS-induced noncancer liver effects, including PPARβ/δ, PPARγ, constitutive androstane receptor (CAR) and pregnane X receptor (PXR), the farnesoid X receptor (FXR), the phosphatidylinositol 3-kinase-serine/threonine kinase Akt (PI3K-Akt) signal transduction pathway, and the nuclear factor kappa B pathway (NF-κB) (Li et al., 2017; Viberg and Eriksson, 2017). Activation of these pathways can be associated with alterations in lipid and glucose metabolism, increased cellular stress, and inflammation (Mackowiak et al., 2018; Li et al., 2017; Mellor et al., 2016; Wang et al., 2014). Thus, the potential involvement and contribution of these different signaling responses to hepatic effects after exposure to the five PFAS will also be considered.

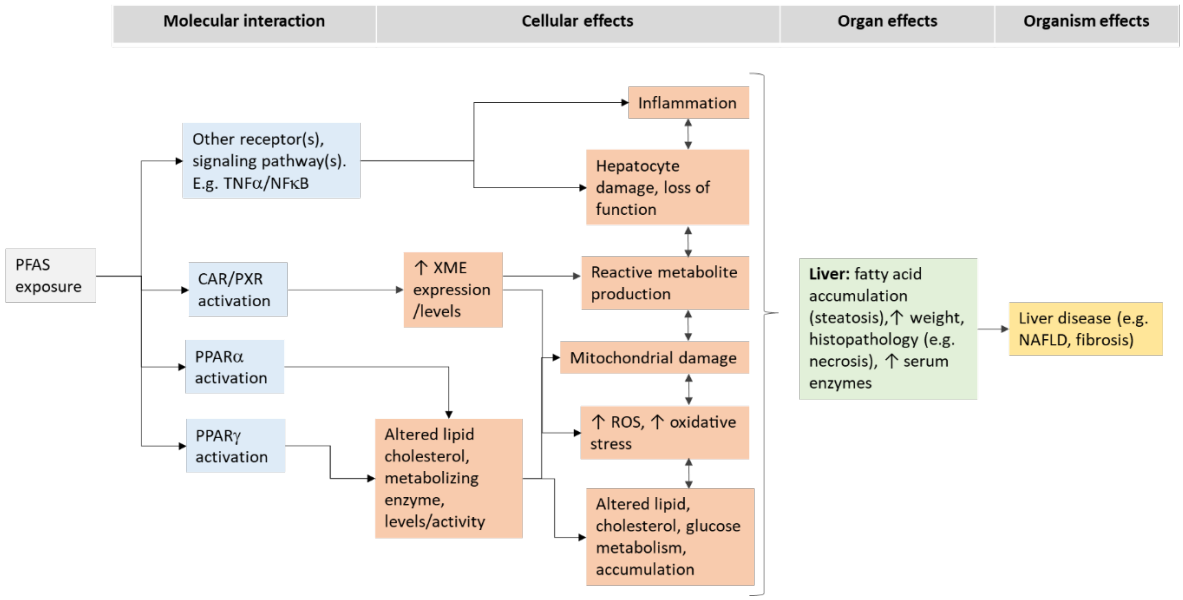


Figure 7. Preliminary proposed mechanistic pathway for per- and polyfluoroalkyl substances (PFAS)-induced noncancer liver effects. Based on previous reviews of perfluorooctane sulfonate (PFOS)- and perfluorooctanoic acid (PFOA)-induced noncancer liver effects in animals (ATSDR, 2018; Li et al., 2017; Viberg and Eriksson, 2017; U.S. EPA, 2016c, d), and proposed adverse outcome pathways for hepatic steatosis (Mellor et al., 2016).

NAFLD = nonalcoholic fatty liver disease; ROS = reactive oxygen species; TNFα = tumor necrosis factor alpha; XME = xenobiotic metabolizing enzymes.

The analysis of the involvement of PPARα and these other signaling cascades in hepatic toxicity after exposure to these five PFAS will focus on the concordance of changes in the specific mechanistic events or separate pathways to effects (i.e., in Figure 7, and as otherwise identified during assessment-specific evaluations) across species to ascertain the relevance of animal studies to human health. The analyses of evidence for each mechanistic event and potential pathway will be qualitatively analyzed for various aspects of the Hill considerations outlined in the EPA Cancer Guidelines framework for MOA analysis (U.S. EPA, 2005a). Given the focus of these analyses, the review will stress the aspects of consistency, coherence, and biological plausibility to ascertain the level of support (or lack thereof), depending on the availability of data. To facilitate this analysis, the following prompting questions and clarifying considerations will be used, depending on the assessment-specific data:

- What is the level of evidentiary support (or lack thereof) for the mechanistic events or signaling pathways, based on the assessment-specific PFAS data? In parallel, are assessment-specific data available to inform the strength of the linkages between events in the pathway or across pathways? Based on the assessment-specific feasibility of doing so (i.e., an adequate database size), answering these questions will incorporate the use of general categories of evidence strength so that the different evaluations can be succinctly summarized. In general, the categories will be based on the following descriptions, each of which will be clarified as support either for or against involvement of an event or pathway: "strong"—independent studies using different experimental models provide consistent and/or coherent evidence; "marginal"—some consistent and/or coherent evidence, although some results may be equivocal or vary from one model to another; and "unclear"—largely inconsistent evidence or insufficient evidence to evaluate.

- Are sufficient assessment-specific data available to inform exposure duration- or level-dependencies for any of the evaluated mechanistic events or pathways?

- Is the assessment-specific evidence (on specific events or pathways in general) consistent with the general biology of the human liver or mechanisms known to be associated with noncancer liver effects in humans? To consider this question, assessments will compare the endpoint-level results across studies on a particular PFAS against the mechanistic understanding/underlying biology for similar effects in the human liver. (Note: this analysis might be informed by studies or reviews on the more robust PFOA/perfluorooctane sulfonate [PFOS] evidence bases.)

- Are responses across studies for these five PFAS assessments indicative of activation of specific mechanisms or signaling pathways conserved across experimental models and designs? To consider this question, assessments will include an evaluation of consistency and coherence across different species and strains of animals, human and animal cell culture models, and *in vivo* humanized animal models, depending on data availability.

- Does the assessment-specific mechanistic information indicate there are likely to be populations or lifestages that may be more susceptible to PFAS-induced liver effects?

The assessment-specific conclusions (and attendant uncertainties) regarding these questions will be used to draw judgments regarding the human relevance of these animal effects, and the rationale for these judgments will be documented transparently within each assessment. As described in EPA guidance (U.S. EPA, 2005a), human relevance is the default and mechanistic evidence will need to be compelling and strong to reach a conclusion otherwise.

9.2.3. Toxicological Relevance of Select Outcomes Observed in Animals

Lastly, the preliminary literature inventory identified the potential for PFAS exposure-mediated effects on several health outcomes for which it is expected to be difficult to identify whether any observed changes (or a lack of changes) are toxicologically relevant, specifically some changes in the kidney and liver (see below). It is expected that in some instances, the synthesis will need to address this issue to inform whether the effects in animals are relevant to interpreting the potential for PFAS exposure to cause a human health effect, and in other instances addressing this issue might be necessary for identifying a level of change for use in determining the potential for adversity or for use in dose-response analysis. It is possible that additional outcomes with similar questions of health relevance might be identified during the development of these assessments. If so, the specifics of the approach selected to address those outcomes will be documented in the assessment(s) and as an update to this protocol. For the aforementioned outcomes, different approaches will be taken, specifically:

1) *Kidney changes in rats, including chronic progressive nephropathy (CPN) and effects that appear to be mediated by an alpha 2u-globulin MOA.* Because the rodent (i.e., male rat)-specific alpha 2u-globulin MOA is not considered relevant to humans, assessments with evidence indicating its involvement will include an evaluation and judgment of the evidence supporting (or not supporting) dependence on this MOA. Specifically, these data will be evaluated against the predefined criteria established by the U.S. EPA (1991a) and/or more recently established criteria, such as those published by Swenberg and Lehman-McKeeman (1999). Relatedly (and possibly overlapping the evaluation of alpha 2u-globulin, because this MOA may exacerbate CPN), there is no human disease analog to the constellation of changes observed in rodent CPN. CPN represents a complex disease process in rats, and its etiology is unknown. Thus, these evaluations will include judgments as to whether all or a subset of the observed changes have adequate evidence to identify dependence on rodent-specific processes, including whether it can be concluded (i.e., based on biological understanding) that the observed kidney endpoints are associated with CPN and the potential for exacerbation of human-relevant disease processes can be ruled out. Data permitting, the assessments will consider whether these conclusions vary by exposure level.

2) *Hepatic changes.* Some individual liver endpoints (and even some constellations of endpoints) might be considered adaptive in nature, possibly leading to the interpretation that some statistically significant changes are not indicative of adverse effects. These endpoints include increased liver weight, cellular hypertrophy, and single cell necrosis/apoptosis. To draw inferences regarding the adversity of these types of liver effects, these assessments will consider the panel recommendations outlined by Hall et al.

(2012) to draw assessment-specific judgments regarding adversity. Briefly, these include evaluation of the available histological data and results suggesting structural degeneration or cellular demise (e.g., apoptosis, oncosis, and/or necrosis), and clinical evidence of hepatocyte damage. As the Hall et al. (2012) recommendations were developed in the context liver tumor formation, consultation of additional reference materials will be considered on an assessment-specific basis. Each assessment will include an explanatory rationale documenting the application of the Hall et al. (2012) recommendations (and any other considerations) to the available evidence.

9.2.4. Other Focused Mechanistic Analyses

Other analyses within the syntheses of mechanistic information will focus on the evidence most useful for informing key uncertainties in the human or animal health effect evidence.

This means that, for example, if extensive and consistent *high*-confidence human or animal evidence is available, the need to synthesize all relevant mechanistic evidence will likely be diminished. In these cases, the analyses will focus on the review and interpretation of smaller sets of mechanistic studies that specifically address controversial or outstanding issues that are anticipated to have a substantial impact on the assessment conclusions. Generally, key uncertainties will be addressed in the mechanistic evidence syntheses by considering the biological understanding of how the effect(s) in question develop or are related. In this way, the analyses can provide information on, for example, (1) potential precursor events when the apical data are uncertain (or unusable for dose-response analyses), (2) the human relevance of animal results when their relevance is unclear or controversial and the human evidence is weak, (3) the shape of the dose-response curve at low exposure levels when this understanding is highly uncertain and data informing this uncertainty are known to exist, or (4) the identification of likely susceptible populations and lifestages. Thus, consideration of biological understanding represents an important component of the evidence analysis. However, mechanistic understanding is not a prerequisite for drawing a conclusion that a chemical causes a given health effect (NTP, 2015; NRC, 2014).

To identify the focused set(s) of studies to use in analyzing critical mechanistic questions other than those outlined in Sections 9.2.1–9.2.3, a stepwise approach will be applied to progressively define the scope of the mechanistic information to be considered throughout assessment development. This stepwise scoping begins during the literature search and screening steps and depends primarily on the potential health hazard signals that arise from the individual human and/or animal health effect studies, or from mechanistic studies that signal potential health hazards that have not been examined in studies of phenotypic, potentially adverse effects. Examples of the focused questions or scenarios triggering these mechanistic evaluations, as well as when during the systematic review they are likely to apply, are listed in Table 17. While the specific methods for evaluating the evidence most relevant to each question will vary, some general considerations for judging the evidence strength in these syntheses are provided below, and if necessary, assessment-specific refinements will be included as updates to the protocol.

1

Table 17. Examples of questions and considerations that can trigger focused analysis and synthesis of mechanistic information

Key assessment-specific uncertainties	Examples of questions and PFAS-specific considerations for identifying the uncertainties and key evidence to analyze
Addressing database completeness based on literature inventories of human, animal, and mechanistic information	• Are there mechanistic studies on an organ system or potential health hazard that were not examined by human or animal studies meeting the PECO criteria? o Depending on the extent of the available data, consider the utility of developing a separate synthesis of evidence versus the utility of a concise, narrative summary (or evidence mapping) to describe these knowledge gaps. Consider whether the mechanistic evidence might be sufficient to substantiate a conclusion on its own (if so, a separate synthesis will be developed).
Addressing questions of inconsistency within the human and animal evidence	• For the health effects of potential concern, is a mechanistic evaluation(s) warranted to inform questions regarding the consistency of the available human or animal studies? Typically, this consideration would focus on health effects that show some indication of an association in epidemiological studies or causality in experimental studies during evidence synthesis. Based on the literature inventory, consider whether mechanistic data are available to inform the specific, key uncertainties in question. Examples of specific scenarios for evaluation include: o If cancer has been observed and tumor types appear to differ across populations (e.g., species or sex), review the literature inventory for mechanistic data that might be relevant to interpreting such differences, and conduct analyses as warranted based on that review. Approaches outlined in the EPA cancer guidelines (U.S. EPA, 2005a) that may be relevant to these analyses will be applied, as appropriate. o If pronounced and unexplained differences in health effect(s)-specific responses are observed across lifestages or populations (e.g., animal strain; human demographic), first consider toxicokinetic differences for the specific PFAS, and then the mechanistic evidence relevant to assessing the potential for health effect-specific biological differences in response (toxicodynamics). Further, inconsistent evidence (i.e., heterogeneous results) across different animal species or human populations might be informed by a review of the evidence relevant to whether different mechanisms may be operant in the different populations (e.g., evidence demonstrating that certain species are more or less sensitive to a certain biological perturbation; evidence that gene polymorphisms are related to variability in response).

Key assessment-specific uncertainties	Examples of questions and PFAS-specific considerations for identifying the uncertainties and key evidence to analyze
Addressing questions of biological plausibility[a] and coherence within the human and animal evidence, and coherence across bodies of evidence	• For the health effects of potential concern, would a mechanistic evaluation(s) of biological plausibility (usually for an individual outcome) or coherence (usually across outcomes) provide meaningful information for interpreting the evidence strength? Typically, this consideration would focus on effects for which the evidence strength for an individual outcome (either for or against an effect) is questionable (e.g., primarily studies of *low* confidence), when a substantial outstanding methodological concern(s) across the relevant studies exists, or when evidence exists for multiple, potentially related (e.g., biologically) outcomes. Based on the literature inventory, consider whether there are mechanistic data available to inform the specific, key uncertainties in question. Examples of specific scenarios for evaluation may include: o If the evidence for a given outcome is weak or uncertain, or when unaddressed methodological concerns identify critical uncertainties in the human or animal findings for a health effect, identify data on mechanistic changes in exposed humans or animals that are likely to be linked to the development or occurrence of the health outcome in question. If enough suitable studies are available, analyze data on changes expected to be related to the phenotypic finding(s) of interest, which can either increase or decrease the evidence strength that the finding(s) is real. It is important to note that the absence of a mechanistic explanation for an association (e.g., the MOA is not understood) will not be used to reduce confidence in observations from human or animal studies. However, the plausibility of an association observed in human or animal studies may be diminished if expected mechanistic findings (e.g., based a known biological dependence) are tested and not apparent. The mechanistic evidence on possible precursors or effects that are known to co-occur with the health outcome of interest are particularly impactful when the changes are observed in the same exposed population presenting the outcome of interest. An understanding of mechanistic pathways (e.g., by identifying and analyzing mechanistic precursor events linked qualitatively or quantitatively to apical health effect[s]; see Section 9.2.2 for additional context) can inform the strength of the evidence integration judgments (see Section 10). o If evidence on multiple health outcomes within an organ system, or possibly across organ systems (e.g., thyroid and nervous system), are available and the strength of the evidence for any single outcome is uncertain, identify biological data that can inform understanding of the relatedness of outcomes within and across systems. Biological understanding or strong mechanistic support (e.g., a shared mechanistic event) for linkages across outcomes can increase the strength of the evidence when changes are related. However, evidence strength may be diminished if an expected pattern among biologically linked outcomes is not observed. Interpretation of the pattern of changes across the outcomes will consider the underlying biological understanding (e.g., one outcome may be expected to precede the other, or be more sensitive). These same considerations inform analyses of the coherence of observed effects across evidence streams during evidence integration (see Section 10.2).

Key assessment-specific uncertainties	Examples of questions and PFAS-specific considerations for identifying the uncertainties and key evidence to analyze
Addressing questions on the human relevance of findings in animals	• For the health effects of potential concern, does the available evidence raise questions of human relevance? Typically, this consideration applies when human evidence is lacking or has results that differ from animal studies, noting that some differences in responses across humans and animals are acknowledged [e.g., for cancer, site concordance is not a requirement for determining the relevance of animal data for humans (U.S. EPA, 2005a); for noncancer nervous system effects, behavioral changes can manifest differently between animals and humans]. The identification of potential differences will also consider ADME information across species, primarily relating to distribution (e.g., to the likely target tissue) and PFAS half-life. Examples of information to identify from the literature inventory, as well as specific scenarios and considerations for these analyses may include: o If there is no evidence indicating that the animal results are irrelevant to humans, summarize existing knowledge on the development of the health effect in each species, including potential differences in PFAS toxicokinetics, and assess the relatedness across species. Note that in the absence of sufficient evidence to the contrary, effects in animal models are assumed to be relevant to humans (ATSDR, 2018; NIEHS, 2015; U.S. EPA, 2005a).[b] o If there is evidence indicating that the mechanisms underlying the effects in animals may not operate in humans, or that the available animal model(s) may not be suitable for the human health outcome(s) of interest, present and analyze the strength of the evidence for and against the human relevance of the observed findings. In addition to considerations specific to the outcome of interest, the analysis will evaluate observations of mechanistic changes in exposed humans for similarities or biological coherence with mechanistic or toxicological changes in experimental animals interpreted to be associated with the health outcome under evaluation. It may also include an evaluation of findings across species known or presumed to be more or less relevant for interpreting potential human toxicity for the health effect(s) in question. In rare instances or for controversial decisions that are likely to drive key assessment conclusions, the analysis may extend to a detailed analysis of a plausible mechanistic pathway(s) or MOA(s) within which each key event and key event relationship is evaluated regarding the likelihood of similarities (e.g., in presence or function) across species. These analyses, regardless of their rigor, will lead to a definitive judgment about whether the animal response is relevant to humans during evidence integration (see Section 10).

Key assessment-specific uncertainties	Examples of questions and PFAS-specific considerations for identifying the uncertainties and key evidence to analyze
Addressing questions on potential susceptibility for hazard identification and dose-response analysis	• For the health effects of potential concern, do the results from the human and animal health effect studies appear to differ by categories that indicate the apparent presence of susceptible populations (e.g., across demographics, species, strains, sexes, or lifestages)? Separately, are there human or animal study data that could identify or clarify population differences in response (e.g., experiments testing sensitivity of responses across lifestages or across genetic variations; observed differences attributable to genetic polymorphisms)? Are there mechanistic data (i.e., based on the literature inventory) that address potential susceptibility factors?[c] If evidence exists for any of these scenarios, information on susceptibility will be reviewed and, if impactful to assessment conclusions, analyzed in detail. Examples of when these analyses are important include: o If the analysis of evidence indicates the likely presence of a sensitive population or lifestage in humans, the groups likely to be at greatest risk will be captured in the evidence integration narrative (see Section 10). In addition, this narrative will discuss whether the appropriate analogous exposures and populations or lifestages were adequately represented or tested in the available human or animal studies, and if not, will identify studies on the most susceptible populations or lifestages as key research needs (see Section 10). o If the analysis of evidence indicates the likely presence of a sensitive population or lifestage in humans, this information will be used to select studies for quantitative analysis (e.g., prioritizing those studies that include such populations [see Section 11]). If specific studies addressing these susceptibilities are unusable for quantitative analysis, susceptibility data may be used to support refined human variability uncertainty factors or probabilistic uncertainty analyses (see Section 11).

Key assessment-specific uncertainties	Examples of questions and PFAS-specific considerations for identifying the uncertainties and key evidence to analyze
Addressing questions on biological understanding to optimize dose-response analysis	• If the human and/or animal health effect data amenable to dose-response analysis are weak[d] or only at high exposure levels, or if the selection of critical parameters for modeling is uncertain, the following analyses will be considered: o When the apical health effect data are highly uncertain or cannot be used with confidence for the purpose of deriving quantitative estimates, mechanistic precursor events linked qualitatively or quantitatively to the phenotypic effect can be evaluated for use as surrogate markers (e.g., based on the strength and completeness of the linkage between mechanistic and phenotypic effects) for deriving quantitative estimates. o When there is a notable lack of understanding of the appropriate exposure metric, biomarker, or modeling parameter for developing quantitative estimates, toxicokinetic and mechanistic understanding of the development of the health effect can inform the most biologically appropriate measure. o When there are dose-response modeling decisions or uncertainties that would be substantially improved by biological or toxicokinetic understanding, mechanistic analyses can improve selection of particular models (e.g., a linear, nonlinear, or threshold model) and help evaluate the appropriateness of integrating/combining data across related outcomes (e.g., based on biological coherence or a conserved MOA). For cancer toxicity values, existing guidance will be consulted (U.S. EPA, 2005a).

[a]As applied herein, biological plausibility describes mechanistic information that either strengthens or weakens an interpretation of the likelihood of an association between exposure and the health effect. The interpretation of biological plausibility considers the existing biological understanding of how the health effect develops and can involve analyses of information at different levels of biological complexity (e.g., molecular, cellular, tissue).

[b]As described in the EPA RfD/RfC Technical Report (2002), "one of the major default assumptions in EPA's risk assessment guidelines is that animal data are relevant for humans [e.g., U.S. EPA (1998), U.S. EPA (1991a), and U.S. EPA (1996b)]. Such defaults are intended to be used in the absence of experimental data that can provide direct information on the relevance of animal data" (U.S. EPA, 2002).

[c]Susceptibility factors include lifestage, demographics and social determinants, behavioral factors, health status, and genetic variability. Although not considered in these analyses, factors that can increase vulnerability, such as other pollutant exposures or differential proximity to exposure sources, are typically considered during a full risk assessment.

[d]Note that "weak" here refers to the study's usability for dose-response analysis specifically. Such studies may be judged to be of *medium* or *high* confidence for the purposes of identifying potential hazards but possess limitations preventing their use for deriving reliable quantitative estimates.

PECO = populations, exposures, comparators, and outcomes; RfC = inhalation reference concentration; RfD = oral reference dose.

If focused areas for additional mechanistic evaluations are identified to inform key assessment-specific uncertainties (e.g., by applying Table 13), the assessments will identify the most impactful studies for evaluation. This could represent only a subset of the potentially relevant studies, particularly if there are many mechanistic studies relevant to the specific question(s).

Because the potential influence of the information provided by the available studies can vary depending on the question(s) or the associated mechanistic events or pathways, the rigor of the analyses will likewise vary from cursory insights drawn from sets of unanalyzed results to detailed evaluations of a subset of the most relevant, individual mechanistic studies. Although the specifics that might be applicable across potential mechanistic topic areas cannot be predefined, the analyses will first consider the studies based on their toxicological relevance to answering the specific question (e.g., model system; specificity of the assay for the effect of interest), potentially refining the focus to a subset of the most relevant studies. This will be particularly important when the set(s) of studies are inconsistent and potentially conflicting, include potentially more informative studies that challenge the necessity of proposed mechanistic relationships between exposure and an apical effect (e.g., altering a receptor-mediated pathway through chemical intervention or using knockout animals), or when it is apparent that particular study design aspects in some or all of the relevant studies are likely to have significant flaws or important uncertainties (e.g., for certain questions, a preliminary review of the exposure methods across the relevant mechanistic studies can flag serious deficiencies). In general, across these assessments, relevant mechanistic information from *in vivo* studies will be prioritized, with preference given to PFAS- and endpoint-relevant exposure routes and exposure designs. Analysis of ex vivo and *in vitro* studies will then be considered, prioritizing those most informative to evaluating the mechanistic events indicated by the *in vivo* data, including studies conducted under conditions most relevant to human exposures and in model systems best replicating *in vivo* human biology.

In some instances, additional literature searches may be warranted, targeting mechanistic events or biological pathways that are not specific to a particular PFAS or group of PFAS. When more rigorous mechanistic analyses are deemed necessary, the review will be aided by the use of pathway-based organizational methods and, if available, established evidence evaluation frameworks. These approaches provide transparency and objectivity for integrating and interpreting the mechanistic events and pathways anchored to the specific questions that have been identified (e.g., anchored to a specific health effect) across diverse sets of relevant data (e.g., human, animal, and *in vitro* studies). The approaches may be facilitated by using organizational tools or frameworks, such as AOPs (see example in Section 9.2.2). As noted above, any additional assessment-specific literature searches and evaluation methods will be described in updates to the protocol.

Based on the analyses and considerations outlined in Sections 9.2.1–9.2.4, the results of the health effect- and assessment-specific mechanistic evidence syntheses will inform both evidence integration and dose-response analyses (see Sections 10 and 11).

10. EVIDENCE INTEGRATION

For the analysis of human health outcomes that might result from chemical exposure, these PFAS assessments will draw integrated judgments across human, animal, and mechanistic evidence for each assessed health effect (see Section 9). As previously discussed in Section 9.2, the approach to evaluating the mechanistic evidence relevant to each assessed health effect will follow a step-wise approach and is expected to vary depending on the nature and impact of the uncertainties identified within each evidence base, as well as the specific mechanistic information available to address those uncertainties. This includes evaluations of mechanistic evidence relevant to the identified key science issues (Section 2.4) prior to or in parallel with evaluations of the phenotypic data in human and animal studies, as well as other focused mechanistic analyses identified during draft development to address key assessment uncertainties (see Section 9.2.4 for a discussion of these scenarios). During evidence integration, a structured and documented process will be used, as follows (and depicted in Figure 8):

- Building from the separate syntheses of the human and animal evidence (see Section 9.1), the strength of the evidence from the available human and animal health effect studies will be summarized in parallel, but separately, using a structured evaluation of an adapted set of considerations first introduced by Sir Bradford Hill (Hill, 1965). Table 19 describes these structured evaluations and the explicit consideration of study confidence within each evaluation domain. Based on the approaches and considerations described in Section 9.2, these summaries will incorporate the relevant mechanistic evidence (or MOA understanding) that informs the biological plausibility and coherence within the available human or animal health effect studies.

- The strength of the animal and human evidence will be considered together in light of inferences across evidence streams. Specifically, the inferences considered during this integration include the human relevance of the animal and mechanistic evidence, coherence across the separate bodies of evidence, and other important information (e.g., judgments regarding susceptibility). Note that without evidence to the contrary, the human relevance of animal findings is assumed.

- A summary judgment is drawn as to whether the available evidence base for each potential human health effect as a whole is sufficient (or insufficient) to indicate that PFAS exposure has the potential to be hazardous to humans[17].

[17] Due to the expected rarity of scenarios where there is "sufficient evidence to judge that a hazard is unlikely" (see description in Table 20 and Section 10.2) and to improve readability, this judgment is not specified in some instances.

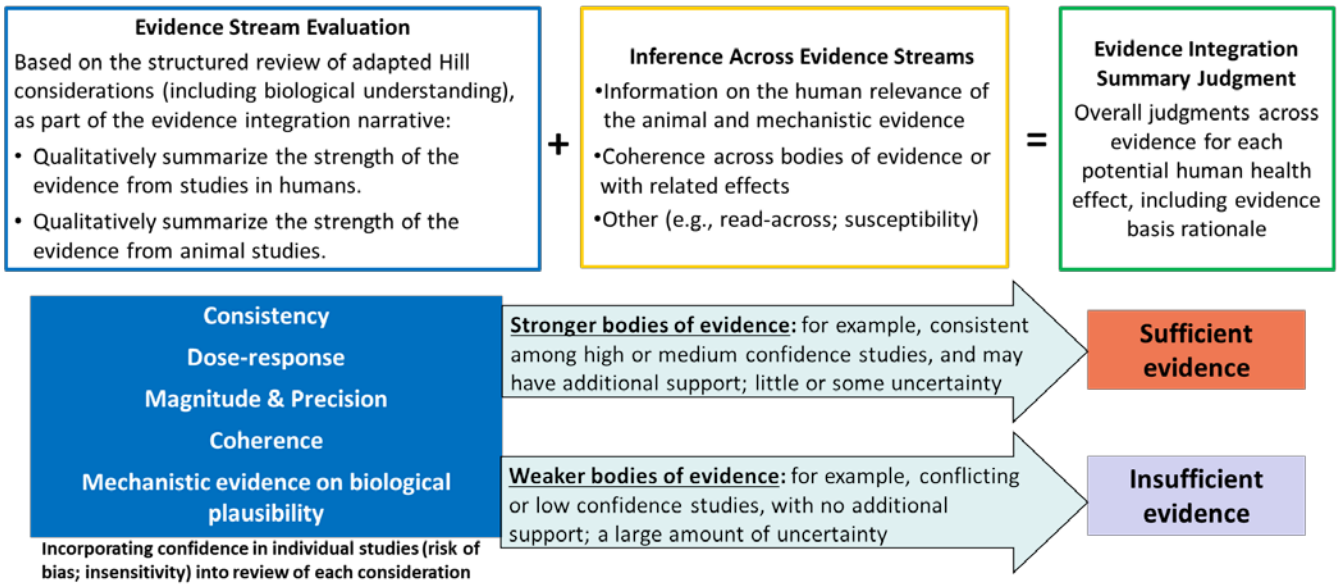


Figure 8. Process for evidence integration. Note that "sufficient evidence" could indicate a judgment of "sufficient evidence for hazard" or "sufficient evidence to judge that a hazard is unlikely", depending on the nature and extent of the available evidence (see **Table 20**).

The decision points within the structured evidence integration process will be summarized in an evidence profile table for each health effect category (see Table 18 for a preliminary template version) in support of the evidence integration narrative. The specific decision frameworks for the structured evaluation of the strength of the human and animal evidence streams and for drawing the overall evidence integration judgment are described in Section 10.1. This process is similar to that used by the Grading of Recommendations Assessment, Development, and Evaluation [GRADE; (Morgan et al., 2016; Guyatt et al., 2011; Schünemann et al., 2011)], which arrives at an overall integration conclusion based on consideration of the body of evidence. As described in Section 9, the human, animal, and mechanistic evidence syntheses serve as inputs providing a foundation for the evidence integration decisions; thus, the major conclusions from these syntheses will be summarized in the evidence profile table (see Table 18 for a preliminary template version) supporting the evidence integration narrative. The evidence profile tables on each potential human health effect evaluated will summarize the judgments and their evidence basis for each step of the structured evidence integration process. Separate sections are included for summarizing the human and animal evidence, for the inference drawn across evidence streams, and for the overall evidence integration judgment. The table presents the key information from the different bodies of evidence that informs each decision.

Table 18. Evidence profile table template

Studies and interpretation	Factors that increase strength	Factors that decrease strength	Summary of evidence streams	Inferences across evidence streams	Overall Evidence Integration Judgment
[Health Effect or Outcome Grouping]					
Evidence in Studies of Humans [may be separated by exposure route[a]]				• Human relevance of findings in animals • Cross-stream coherence • Other inferences: o Information on susceptibility o MOA analysis inferences o Relevant information from other sources (e.g., read across)	Describe judgment regarding whether there is sufficient (or insufficient) evidence to identify a potential human health hazard, integrating evidence across streams and including a summary of the models and range of dose levels upon which the judgment is primarily reliant.
• References • Study confidence • Study design description	• Consistency • Dose-response gradient • Coherence of observed effects • Effect size • Mechanistic evidence providing plausibility • *Medium* or *high* confidence studies[b]	• Unexplained inconsistency • Imprecision • *Low* confidence studies[b] • Evidence demonstrating implausibility	Qualitative summary of the strength of the evidence from human studies based on the factors at left, including the primary evidence basis and considering: • Results across human epidemiological and controlled exposure studies • Human mechanistic evidence informing biological plausibility (e.g., precursor events linked to adverse outcomes)		
Evidence from Animal Studies [may be separated by exposure route[a]]					
• References • Study confidence • Study design description	• Consistency and/or Replication • Dose-response gradient • Coherence of observed effects • Effect size • Mechanistic evidence providing plausibility • *Medium* or *high* confidence studies[b]	• Unexplained inconsistency • Imprecision • *Low* confidence studies[b] • Evidence demonstrating implausibility	Qualitative summary of the strength of the evidence for an effect in animals based on the factors at left, including the primary evidence basis and considering: • Results across animal toxicological studies • Animal mechanistic evidence informing biological plausibility (e.g., precursor events linked to adverse outcomes)		

[a]In addition to exposure route, the summaries of the strength of each evidence stream may include multiple rows- e.g., by study confidence, population, or species, if this informed the analysis of results heterogeneity.

[b]Study confidence, based on evaluation of risk of bias and study sensitivity (see Section 6), and information on susceptibility will be considered when evaluating each of the other factors that increase or decrease strength (e.g., consistency). Notably, lack of findings in studies deemed insensitive neither increases or decreases strength.

10.1. EVALUATING THE STRENGTH OF THE HUMAN AND ANIMAL EVIDENCE STREAMS

As summarized above, prior to drawing overall evidence integration judgments about whether exposure to one of these five PFAS has the potential to cause certain health effect(s) in humans given relevant exposure circumstances, the strength of evidence for the available human and animal evidence will be evaluated and summarized. For each assessed health effect or health effect grouping (see Section 5 for examples of the endpoints that will be considered within each health effect category), the relevant mechanistic evidence in exposed humans and animals (or in their cells, relevant NAMs or in silico models), which will be synthesized based on the approaches and considerations in Section 9.2, will be integrated with the evidence from the available studies of phenotypic effects in humans and animals. The considerations previously outlined in Table 15 (the different features of the evidence considered and summarized during evidence synthesis; see Section 9) will be evaluated by the specific PFAS assessment teams within the context of how they affect judgments of the strength of evidence (see Table 19), which will directly inform the overall evidence integration judgment (see Section 10.2). The evaluation of the strength of the human or animal health effects evidence (i.e., based on the considerations in Table 19) will preferably occur at the most specific health outcome level possible (e.g., an analysis at the level of decreased pulmonary function is generally preferable to an analysis of respiratory system effects), if there is an adequate set of studies for analyses at this level and considering the interrelatedness of the available outcomes. If studies on a target system are sparse or varied, or if the interpretation of evidence strength relies largely on the consideration of coherence across related outcomes, then the analyses may need to be conducted at a broader health effect level. The factors judged to increase or decrease the strength of the evidence will be summarized in tabular format using the evidence profile table template in Table 18 to transparently convey expert judgments made throughout the evidence synthesis and integration processes. The evidence profile table allows for consistent documentation of the supporting rationale for each decision.

Table 19. Considerations that inform evaluations of the strength of the human and animal evidence

Consideration	Increased evidence strength (of the human or animal evidence)	Decreased evidence strength (of the human or animal evidence)
The structured categories and criteria inTable 20 (Section 10.2) will guide the application of strength-of-evidence judgments for an outcome or health effect. Evidence synthesis scenarios that do not warrant an increase or decrease in evidence strength for a given consideration will be considered "neutral" and are not described in this table (and, in general, will not be captured in the assessment-specific evidence profile tables).		
Risk of bias; sensitivity (across studies)	• An evidence base of *high*- or *medium*-confidence studies increases strength.	• An evidence base of mostly *low*-confidence studies decreases strength. An exception to this is an evidence base of studies in which the primary issues resulting in *low* confidence are related to insensitivity. This may increase evidence strength in cases where an association is identified because the expected impact of study insensitivity is towards the null. • Decisions to increase strength for other considerations in this table should generally not be made if there are serious concerns for risk of bias.
Consistency	• Similarity of findings for a given outcome (e.g., of a similar magnitude, direction) across independent studies or experiments increases strength, particularly when consistency is observed across populations (e.g., geographical location) or exposure scenarios in human studies, and across laboratories, populations (e.g., species), or exposure scenarios (e.g., duration; route; timing) in animal studies.	• Unexplained inconsistency (i.e., conflicting evidence; see U.S. EPA (2005a)) decreases strength. Generally, strength should not be decreased if discrepant findings can be reasonably explained by study confidence conclusions; variation in population or species, sex, or lifestage; exposure patterns (e.g., intermittent or continuous); exposure levels (low or high); or exposure duration.

Consideration	Increased evidence strength (of the human or animal evidence)	Decreased evidence strength (of the human or animal evidence)
Strength (effect magnitude) and precision	• Evidence of a large magnitude effect (considered either within or across studies) can increase strength. Effects of a concerning rarity or severity can also increase strength, even if they are of a small magnitude. • Precise results from individual studies or across the set of studies increases strength, noting that biological significance is prioritized over statistical significance.	• Strength may be decreased if effect sizes that are small in magnitude are concluded not to be biologically significant, or if there are only a few studies with imprecise results.

Consideration	Increased evidence strength (of the human or animal evidence)	Decreased evidence strength (of the human or animal evidence)
Biological gradient/dose-response	• Evidence of dose-response increases strength. Dose-response may be demonstrated across studies or within studies and it can be dose- or duration-dependent. It also may not be a monotonic dose-response (monotonicity should not necessarily be expected, e.g., different outcomes may be expected at low vs. high doses due to activation of different mechanistic pathways or induction of systemic toxicity at very high doses). • Decreases in a response after cessation of exposure (e.g., symptoms of current asthma) also may increase strength by increasing certainty in a relationship between exposure and outcome (this is most applicable to epidemiology studies because of their observational nature).	• A lack of dose-response when expected based on biological understanding and having a wide range of doses/exposures evaluated in the evidence base can decrease strength. • In experimental studies, strength may be decreased when effects resolve under certain experimental conditions (e.g., rapid reversibility after removal of exposure). However, many reversible effects are of high concern. Deciding between these situations is informed by factors such as the toxicokinetics of the chemical and the conditions of exposure [see U.S. EPA (1998)], endpoint severity, judgments regarding the potential for delayed or secondary effects, as well as the exposure context focus of the assessment (e.g., addressing intermittent or short-term exposures). • In rare cases, and typically only in toxicology studies, the magnitude of effects at a given exposure level might decrease with longer exposures (e.g., due to tolerance or acclimation). Like the discussion of reversibility above, a decision about whether this decreases evidence strength depends on the exposure context focus of the assessment and other factors. • If the data are not adequate to evaluate a dose-response pattern, then strength is neither increased or decreased.

Consideration	Increased evidence strength (of the human or animal evidence)	Decreased evidence strength (of the human or animal evidence)
Coherence	• Biologically related findings within an organ system, or across populations (e.g., sex) increase strength, particularly when a temporal- or dose-dependent progression of related effects is observed within or across studies, or when related findings of increasing severity are observed with increasing exposure.	• An observed lack of expected coherent changes (e.g., well-established biological relationships) will typically decrease evidence strength. However, the biological relationships between the endpoints being compared and the sensitivity and specificity of the measures used need to be carefully examined. The decision to decrease evidence strength depends on the availability of evidence across multiple related endpoints for which changes would be anticipated, and it considers factors (e.g., dose and duration of exposure, strength of expected relationship) across the studies of related changes.
Mechanistic evidence related to biological plausibility	• Mechanistic evidence of precursors or health effect biomarkers in well-conducted studies of exposed humans or animals, in appropriately exposed human or animal cells, or other relevant human, animal, or in silico models (including new approach methods, NAMs) increases strength, particularly when this evidence is observed in the same cohort/population exhibiting the phenotypic health outcome. • Evidence of changes in biological pathways or support for a proposed MOA in appropriate models also increases strength, particularly when support is provided for rate-limiting or key events, or is conserved across multiple components of the pathway or MOA.	• Mechanistic understanding is not a prerequisite for drawing a conclusion that a chemical causes a given health effect; thus, an absence of knowledge should not be used a basis for decreasing strength (NTP, 2015; NRC, 2014). • Mechanistic evidence in well-conducted studies (see examples of evidence types at left) that demonstrates that the health effect(s) are unlikely to occur, or only likely to occur under certain scenarios (e.g., above certain exposure levels), can decrease evidence strength. A decision to decrease strength depends on an evaluation of the strength of the mechanistic evidence supporting vs. opposing biological plausibility, as well as the strength of the health effect-specific findings (e.g., stronger health effect data require more certainty in mechanistic evidence opposing plausibility).

1

For human and animal evidence, the analyses of each consideration in Table 19 will be used to qualitatively summarize the strength-of-evidence for the separate evidence streams in the evidence integration narrative. Table 20 provides the criteria that will guide how to draw the judgment for each health effect, and the terms that will be used to summarize those evidence integration judgments.

10.2. OVERALL EVIDENCE INTEGRATION JUDGMENTS

Evidence integration combines decisions regarding the strength of the animal and human evidence with considerations regarding mechanistic information on the human relevance of the animal evidence, relevance of the mechanistic evidence to humans (especially in cases where animal evidence is lacking), coherence across bodies of evidence, and information on susceptible populations and lifestages, based on the considerations and analyses outlined in Section 9.2. This evidence integration decision process will culminate in an evidence integration narrative that summarizes the judgments regarding the evidence for each potential health effect (i.e., each noncancer health effect and specific type of cancer, or broader grouping of related outcomes). For each health effect, this narrative will include:

- A descriptive summary of the primary judgments about the evidence informing the potential for health effects in exposed humans, based on the following analyses:
    - evaluations of the strength of the available human and animal evidence (see Section 10.1);
    - consideration of the coherence of findings (i.e., the extent to which the evidence for health effects and relevant mechanistic changes are similar) across human and animal studies;
    - other information on the human relevance of findings in animals (see Section 9.2); and
    - conclusions drawn based on the predefined mechanistic analyses (see Sections 9.2.1–9.2.3), as well as those based on analyses identified during stepwise consideration of the health effect-specific evidence during draft development (see Section 9.2.4).
- A summary of key evidence supporting these judgments, highlighting the evidence that was the primary driver of these judgments and any notable issues (e.g., data quality; coherence of the results), and a narrative expression of confidence (a summary of strengths and remaining uncertainties) for these judgments.
- Information on the general conditions of expression of these health effects (e.g., exposure routes and levels in the studies that were the primary drivers of these judgments), noting that these conditions will be clarified during dose-response analysis (see Section 11).

- Indications of potentially susceptible populations or lifestages (i.e., an integrated summary of the available evidence on potential susceptible populations and lifestages drawn across the syntheses of the human, animal, and mechanistic evidence)[18].

- A summary of key assumptions used in the analysis, which are generally based on EPA guidelines and which are largely captured in this protocol.

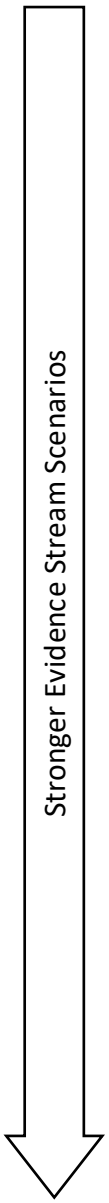
- Strengths and limitations of the evidence integration judgments, including key uncertainties and data gaps, as well as the limitations of the systematic review. As noted in Section 4.2.2, for one or more of these five PFAS assessments, characterization of the uncertainties in the animal evidence is expected to include a discussion of the reliance on short-term oral exposure studies in rats. Similarly, the characterization of uncertainty in the human evidence is expected to include a discussion of potential confounding by PFAS other than the PFAS of interest.

In short, the evidence integration narrative will present a qualitative summary of the strength of each evidence stream and an overall judgment across all relevant evidence, with exposure context provided. For each health effect or specific cancer type of potential concern, the first sentence of the evidence integration narrative will include the summary judgment [see description below for how these judgments help inform selection of a descriptor for carcinogenicity (U.S. EPA, 2005a)]. Assessments will also include an evidence profile table (see Table 18) to support the evidence integration narrative by providing the major decisions and supporting rationale. Table 20 describes the categories of evidence integration judgments that will be used in these PFAS assessments and provides examples of database scenarios that fit each category of evidence. These summary judgments provide a succinct and clear representation of the decisions from the more detailed analyses of whether (or not) the evidence strength indicates that PFAS exposure has the potential to cause the human health effect(s) under the necessary conditions of exposure. Consistent with EPA non-cancer and cancer guidelines, a judgment that the evidence supports an apparent lack of an effect of PFAS exposure on the health effect(s) will only be used when the available data are considered robust for deciding that there is no basis for human hazard concern; lesser levels of evidence suggesting a lack of an effect will be characterized as "insufficient."

[18]One or more of these five PFAS assessments may include consideration of information outside of their PFAS-specific database to address this aspect of the evidence integration narrative. These PFAS-specific data gaps and uncertainties appear to extend beyond poorly studied health effects, and the discussion of missing information on potential populations, sexes, or lifestages that are likely to be more susceptible to developing a specific health effect may consider information from reviews of other PFAS.

Table 20. Evidence integration judgments for characterizing potential human health hazards in the evidence integration narrative

Evidence Integration Judgment[1]	Evidence in Studies of Humans		Evidence in Animal Studies	Inferences Across Evidence Streams
Sufficient evidence for hazard	A judgment of ***sufficient evidence for hazard*** requires that a scenario below is met for either the *evidence in studies of humans* OR *evidence in animal studies*, incorporating the considerations outlined under *inferences across evidence streams*. The scenarios justifying this judgment span a broad range of overall evidence strength and examples are provided below, starting with the weakest evidence [2].			
	• Strong mechanistic evidence in well-conducted studies of exposed humans (*medium* or *high* confidence) or human cells (including NAMs), in the absence of other substantive data, where an informed evaluation has determined that the data are reliable for assessing toxicity relevant to humans and the mechanistic events have been causally linked to the development of the health effect of interest[3]. • A single *high* or *medium* confidence study without supporting coherent evidence or concern for unexplained inconsistency. Specifically, there are no comparable studies of similar confidence and sensitivity providing conflicting evidence, or the differences can be reasonably explained by, e.g., the populations or exposure levels studied (U.S. EPA, 2005a). • Multiple studies showing generally consistent findings, including at least one *high* or *medium* confidence study and supporting evidence, but with some residual uncertainty due to potential chance, bias, or confounding (e.g., effect estimates of low magnitude or small effect sizes given what is known about the endpoint; uninterpretable	Stronger Evidence Stream Scenarios	• Strong mechanistic evidence in well-conducted studies of animals or animal cells (including NAMs), in the absence of other substantive data, where an informed evaluation has determined the assays are reliable for assessing toxicity relevant to humans and the mechanistic events have been causally linked to the development of the health effect [3]. • A single *high* or *medium* confidence experiment in the absence of comparable experiment(s) of similar confidence and sensitivity providing conflicting evidence [4] (evidence that cannot be reasonably explained, e.g., by respective study designs or differences in animal model; (U.S. EPA, 2005a). • At least one *high* or *medium* confidence study with supporting information increasing the strength of the evidence. Although the results are largely consistent, notable uncertainties remain. However, in scenarios when inconsistent evidence or evidence indicating nonspecific effects exist, it is not judged to reduce or discount the level of concern regarding the positive findings, or it is not	• Supplemental evidence (e.g., structure-activity data; chemical class information; other NAMs) is judged to increase the strength of limited or near-equivocal, chemical-specific human or animal evidence to ***sufficient evidence for hazard***. • Coherent or biologically consistent findings across evidence streams increases the strength of limited or near-equivocal human or animal evidence (e.g., single or few *high* or *medium* confidence studies with some conflicting evidence) to ***sufficient evidence for hazard***. • The strength of the evidence is decreased because mechanistic information (even if it does not provide MOA understanding) raises uncertainties regarding the human and/or animal evidence, but overall the evidence is still considered strong enough to result in a judgment of ***sufficient evidence for hazard.*** • The strength of the evidence is decreased because findings across

	patterns with respect to exposure levels). Alternatively, a single *high* or *medium* confidence study with a large magnitude or severity of the effect, a dose-response gradient, or other factors that increase the evidence strength, without serious residual uncertainties. In both scenarios, associations with related endpoints, including mechanistic evidence from exposed humans, can address uncertainties relating to exposure response, temporality, coherence, and biological plausibility, and any conflicting evidence is not from a comparable body of higher confidence, sensitive studies[4]. • A set of *high* or *medium* confidence independent studies reporting an association between the exposure and the health outcome, with reasonable confidence that alternative explanations, including chance, bias, and confounding, can be ruled out across studies. The set of studies is primarily consistent, with reasonable explanations when results differ; and an exposure response gradient is demonstrated. Supporting evidence, such as associations with biologically related endpoints in human studies (coherence) or large estimates of risk or severity of the response, may help to rule out alternative explanations. Similarly, mechanistic evidence from exposed humans may serve to address uncertainties relating to exposure-response, temporality, coherence, and biological plausibility (i.e., providing evidence consistent with an explanation for how exposure could cause the health effect		from a comparable body of higher confidence, sensitive studies [4]. The additional support provided includes either consistent effects across laboratories or species; coherent effects across multiple related endpoints; an unusual magnitude of effect, rarity, age at onset, or severity; a strong dose-response relationship; or consistent observations across exposure scenarios (e.g., route, timing, duration), sexes, or animal strains. Mechanistic evidence in animals may serve to provide this support or otherwise address residual uncertainties. • A set of *high* or *medium* confidence experiments with consistent findings of adverse or toxicologically significant effects across multiple laboratories, exposure routes, experimental designs (e.g., a subchronic study and a two-generation study), or species; and the experiments reasonably rule out the potential for nonspecific effects to have caused the effects of interest. Any inconsistent evidence (evidence that cannot be reasonably explained based on study design or differences in animal model) is from a set of experiments of lower confidence or sensitivity. To reasonably rule out alternative explanations, multiple additional factors in the set of experiments exist, such as: coherent effects across biologically related endpoints; an unusual magnitude of effect, rarity, age at onset, or severity; a strong dose-response relationship; or consistent observations across animal lifestages, sexes, or strains. Similarly,	evidence streams are conflicting (U.S. EPA, 2005a) or biologically inconsistent, but a judgment of ***sufficient evidence for hazard*** is supported by review of the adversity and human relevance (prioritizing findings relevant to human toxicity) of the effects. • The strength of the evidence is neither increased or decreased due to a lack of experimental information on the human relevance of the animal evidence or mechanistic understanding (mechanistic evidence may exist, but it is inconclusive); in these cases, the animal data are judged not to conflict with current biological understanding and thus are assumed to be relevant, while findings in humans and animals are presumed to be real unless proven otherwise. • For the strongest animal evidence, there is mechanistic understanding that the findings are expected to occur and progress in humans. Most notably, an MOA interpreted with reasonable certainty would rule out alternative explanations. • For the strongest evidence, there is adequate testing of potentially susceptible lifestages and populations, based on the effect(s) of interest and chemical knowledge (e.g., toxicokinetics).

	based on current biological knowledge).		mechanistic evidence (e.g., precursor events linked to adverse outcomes) in animal models may exist to address uncertainties in the evidence base.	
Insufficient evidence	A judgment of ***insufficient evidence*** requires that a scenario below is met for <u>both</u> the *evidence in studies of humans* AND *evidence in animal studies*, incorporating the considerations outlined under *inferences across evidence streams*.			
	• A body of evidence, including scenarios with one or more *high* or *medium* confidence studies reporting an association between exposure and the health outcome, where either (1) conflicting evidence exists in studies of similar confidence and sensitivity [4,5] OR (2) considerable methodological uncertainties remain across the body of evidence (typically related to exposure or outcome ascertainment, including temporality), AND there is no supporting coherent evidence that increases the overall evidence strength. • A set of only *low* confidence studies. • No studies of exposed humans or well-conducted studies of human cells. • A set of largely null studies that does not meet a scenario for ***sufficient evidence to judge that a hazard is unlikely***.	• A body of evidence, including scenarios with one or more *high* or *medium* confidence experiments reporting effects but without supporting coherent evidence that increases the overall evidence strength, where conflicting evidence exists from a set of sensitive experiments of similar or higher confidence (can include mechanistic evidence) [4,5]. • A set of only *low* confidence experiments. • No animal studies or well-conducted studies of animal cells. • The available endpoints are not informative to the hazard question under evaluation. • A set of largely null studies that does not meet the criteria for ***sufficient evidence to judge that a hazard is unlikely***.	• The *evidence in animal studies* meets a scenario for ***sufficient evidence for hazard***, but strong experimental evidence (e.g., a MOA interpreted with reasonable certainty) indicates the findings in animals are unlikely to be relevant to humans. • The evidence meets a scenario for ***sufficient evidence to judge that a hazard is unlikely***, but there is inadequate testing of susceptible populations and lifestages, or data conflict across evidence streams. • The *evidence in animal studies* meets a scenario for ***sufficient evidence to judge that a hazard is unlikely***, but the database lacks experimental support that the models are relevant to humans for the effect of interest.
Sufficient evidence to judge that a hazard is unlikely [6]	A judgment of ***sufficient evidence to judge that a hazard is unlikely*** requires that a scenario below is met for <u>either</u> the *evidence in studies of humans* OR *evidence in animal studies*, incorporating the considerations outlined under *inferences across evidence streams*.			
	• Several *high* confidence studies showing null results (for example, an odds ratio of 1.0), ruling out alternative explanations including chance, bias, and confounding with reasonable confidence. Each of the studies should have used an optimal outcome and	• A set of *high* confidence experiments examining a reasonable spectrum of endpoints relevant to a type of toxicity that demonstrate a lack of biologically significant effects across multiple species, both sexes (if applicable), and a broad range of exposure levels. The data are compelling in that the experiments have	• There is adequate testing of susceptible populations and lifestages. • When the *evidence in animal studies* meets a scenario for this	

	exposure assessment and adequate sample size (specifically for higher exposure groups and for susceptible populations). The overall set should include the full range of levels of exposures that human beings are known to encounter, and an evaluation of an exposure-response gradient.	examined the range of scenarios across which health effects in animals could be observed, and an alternative explanation (e.g., inadequately controlled features of the studies' experimental designs; inadequate sample sizes) for the observed lack of effects is not available. The experiments were designed to specifically test for the effects of interest, including suitable exposure timing and duration, post- exposure latency, and endpoint evaluation procedures.	judgment, there is experimental support that the models are relevant to humans for the effect of interest and no conflicting human evidence exists. • When the *evidence in studies of humans* meets a scenario for this judgment and conflicting animal data exist, mechanistic information indicates that the animal data are unlikely to be relevant to humans. • When multiple *high* confidence animal experiments and studies in humans indicate lack of an effect, but the evidence does not meet a scenario for ***sufficient evidence to judge that a hazard is unlikely***, strong mechanistic evidence in models relevant to humans supports lack of an effect such that the totality of evidence supports this judgment.

[1] These categories are based on those indicated for use in hazard characterization from the existing EPA guidelines for noncancer health effects (i.e., (U.S. EPA, 1996a, 1991b, 1998)) and, as described in those guidance documents, they depend heavily on expert judgment (note: as applied herein, the process of 'evidence integration' is synonymous with 'weight of evidence'). The evidence integration judgment for each assessed health effect will be included as part of an evidence integration narrative, the specific documentation of the various expert decisions and evidence-based (or default) rationales are summarized in an evidence profile table, and the judgement will be contextualized based on the primary supporting evidence (experimental model or observed population, and exposure levels tested or estimated). Importantly, as discussed in Section 10.1, these judgments may be based on analyses of grouped outcomes at different levels of granularity (e.g., motor activity versus neurobehavioral effects versus nervous system effects) depending on the specifics of the health effect evidence base. Health effects characterized as having ***sufficient evidence for hazard*** will be evaluated for use in dose-response assessment.

[2] Qualitative descriptions of differences in the strength of the evidence across different health effects judged as having ***sufficient evidence for hazard*** are useful for other assessment decisions, including prioritizing outcomes in quantitative analyses and characterizing assessment uncertainties. Thus, for all evidence scenarios, but particularly for those in the lower end of this range, it is important to characterize the uncertainties in the evidence base within the evidence integration narrative and convey the evidence strength to subsequent steps, including toxicity values developed based on those effects. Existing guidance defines the minimum evidence necessary to judge that a health hazard could exist as one adverse endpoint from one well-conducted study U.S. EPA (1998); this has been expanded in this table to better incorporate mechanistic evidence, including new approach methods (NAMs) data.

[3] Scientific understanding of toxicity mechanisms and of the human implications of new toxicity testing methods (e.g., from high-throughput screening, from short-term in vivo testing of alternative species, or from new in vitro and in silico testing and other NAMs) will continue to increase. Thus, the sufficiency of mechanistic evidence alone for identifying potential human health hazards is expected to increase as the science evolves. The decision to identify a potential human hazard based on these data is an expert judgment dependent on the state-of-the-science at the time of review.

[4] Scenarios with unexplained heterogeneity across sets of studies with similar confidence and sensitivity can be considered either ***sufficient evidence for hazard*** or ***insufficient evidence***, depending on the expert judgment of the overall weight of evidence. Specifically, this judgment considers the level of support (or lack thereof) provided by evaluations of the magnitude or severity of the effects, coherence of related findings (including mechanistic evidence), dose-response, and biological plausibility, as well as the comparability of the supporting and conflicting evidence (e.g., the specific endpoints tested, or the methods used to test them; the specific sources of bias or insensitivity in the respective sets of studies). The evidence-specific factors supporting either evidence integration judgment will be clearly articulated in the evidence integration narrative.

[5] When the database includes at least one well-conducted study and a hazard characterization judgment of ***insufficient evidence*** is drawn, quantitative analyses may still be useful for some purposes (e.g., providing a sense of the magnitude and uncertainty of estimates for health effects of potential concern, ranking potential hazards, or setting research priorities), but not for others (see related discussions in (U.S. EPA, 2005a)). It is critical to transparently convey the extreme uncertainty in any such estimates.

[6] The criteria for this category are intentionally more stringent than those justifying a conclusion of ***sufficient evidence for hazard***, consistent with the "difficulty of proving a negative" (as discussed in (U.S. EPA, 1996a, 1991b, 1998)).

1

For evaluations of carcinogenicity, consistent with EPA's Cancer Guidelines (U.S. EPA, 2005a), one of EPA's standardized cancer descriptors will be used as a shorthand characterization of the evidence integration narrative, describing the overall potential for human carcinogenicity across all potential cancer types. These are: (1) *carcinogenic to humans*, (2) *likely to be carcinogenic to humans*, (3) *suggestive evidence of carcinogenic potential*, (4) i*nadequate information to assess carcinogenic potential*, or (5) *not likely to be carcinogenic to humans*. More than one descriptor can be used when a chemical's effects differ by exposure level or route (U.S. EPA, 2005a); if the database supports such an analysis, these decisions will be clarified based on a more thorough review of the mechanistic evidence or more detailed dose-response analysis (see Section 11). In some cases, mutagenicity will also be evaluated (e.g., when there is evidence of carcinogenicity), because it influences the approach to dose-response assessment and subsequent application of adjustment factors for exposures early in life (U.S. EPA, 2005a, b).

An appropriate cancer descriptor will be selected as described in EPA Cancer Guidelines (U.S. EPA, 2005a). For each cancer subtype, an evidence integration narrative and summary judgment will be provided, as described above. The cancer descriptor will consider the interrelatedness of cancer types potentially related to PFAS exposure, consistency across the human and animal evidence for any cancer type [noting that site concordance is not required (U.S. EPA, 2005a)], and the uncertainties associated with each assessment-specific conclusion. In general, however, if a systematic review of more than one cancer type was conducted, then the overall judgment and discussion of evidence strength in the evidence integration narrative for the cancer type(s) with the strongest evidence for hazard will be used to inform selection of the cancer descriptor, with each assessment providing a transparent description of the decision rationale. The cancer descriptor and evidence integration narrative for potential carcinogenicity, including application of the MOA framework, will consider the conditions of carcinogenicity, including exposure (e.g., route; level) and susceptibility (e.g., genetics; lifestage), as the data allow (Farland, 2005; U.S. EPA, 2005a, b).

10.3. HAZARD CONSIDERATIONS FOR DOSE-RESPONSE

This section outlines how these assessments will consider and describe the transition from hazard identification to dose-response analysis, highlighting (1) information that will inform the selection of outcomes or broader health effect categories for which toxicity values will be derived, (2) whether toxicity values can be derived to protect specific populations or lifestages, (3) how dose-response modeling will be informed by toxicokinetic information, and (4) information aiding the identification of biologically based benchmark response (BMR) levels. The pool of outcomes and study-specific endpoints will be discussed to identify which categories of effects and study designs are considered the strongest and most appropriate for quantitative assessment of a given health effect. Health effects that were analyzed in human studies in relation to exposure levels within or closer to the range of exposures encountered in the environment will be considered

particularly informative, as are animal studies testing a broad range of exposure levels and including levels in the lower dose region. When there are multiple endpoints for an organ/system, considerations for characterizing the overall impact on this organ/system will be discussed, including the severity and longevity of the effects. For example, if there are multiple histopathological alterations relevant to liver function changes, liver necrosis may be selected as the most representative endpoint to consider for dose-response analysis. This section may review or clarify which endpoints or combination of endpoints in each organ/system characterize the overall effect for dose-response analysis. For cancer types, consideration will be given to the overall risk of multiple types of tumors. Multiple tumor types (if applicable) will be discussed, and a rationale given for any grouping.

Biological considerations that are important for dose-response analysis (e.g., that could help with selection of a BMR) will be discussed. The impact of route of exposure on toxicity to different organs/systems will be examined, if appropriate. The existence and validity of PBPK models or toxicokinetic information that may allow the estimation of internal dose for route-to-route extrapolation will be presented (see additional discussion and decision points in Section 11.2). In addition, mechanistic evidence analyses that will influence the dose-response analyses will be highlighted (see Section 9.2 for specific considerations), for example, evidence related to susceptibility or potential shape of the dose-response curve.

This section will also describe the evidence regarding populations and lifestages that appear to be susceptible to the health hazards identified and factors that are likely to increase the risk of developing (or exacerbating) these health effects, depending on the available evidence. This section will include this discussion even if there are no specific data on the effects of exposure to the PFAS of interest in the potentially susceptible population. Table 16 in Section 9 outlines some of the specific factors that will be considered for discussion and summaries of the evidence with respect to patterns across studies pertinent to consistency, coherence, and the magnitude and direction of effect measures. At a minimum, consideration will be given to discussion of information relevant to infants and children, pregnant women, and women of childbearing age.

The section will consider options for using susceptible population data in the dose-response analysis. In particular, an attempt will be made to highlight where it might be possible to develop separate risk estimates for a specific population or lifestage or to determine whether evidence is available to select a data-derived uncertainty factor.

11. DOSE-RESPONSE ASSESSMENT: SELECTING STUDIES AND QUANTITATIVE ANALYSIS

The previous sections of this protocol describe how systematic review principles will be applied to evaluate studies (for potential bias and sensitivity) and reach evidence integration conclusions on potential human health effects associated with exposure to the PFAS of interest. Selection of specific data sets for dose-response assessment and performance of the dose-response assessment will be conducted after hazard identification is complete and involves database- and chemical-specific biological judgments that build from decisions made at earlier stages of assessment development. A number of Environmental Protection Agency (EPA) guidance and support documents detail data requirements and other considerations for dose-response modeling, especially EPA's *Benchmark Dose Technical Guidance* (U.S. EPA, 2012), EPA's *Review of the Reference Dose and Reference Concentration Processes* (U.S. EPA, 2002), *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005a), and *Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens* (U.S. EPA, 2005b). This section of the protocol provides an overview of considerations for conducting the dose-response assessment, particularly statistical considerations specific to dose-response analysis that support quantitative risk assessment. Importantly, these considerations do not supersede existing EPA guidance.

Dose-response assessments will be performed for both noncancer and cancer health hazards, and for both oral and inhalation routes of exposure following exposure[19] to the chemical of interest, if supported by existing data. For noncancer hazards, an oral reference dose (RfD) and/or an inhalation reference concentration (RfC) will be derived when possible. An RfD or an RfC is an estimate, with uncertainty spanning perhaps an order of magnitude, of an exposure to the human population (including susceptible subgroups) that is likely to be without an appreciable risk of deleterious health effects over a lifetime (U.S. EPA, 2002). These health effects may also include cancer effects [e.g., in a case where a nonlinear MOA is concluded that indicates a key precursor event necessary for carcinogenicity does not occur below a specific exposure level (U.S. EPA, 2005a); see Section 11.2.3]. Reference values are not predictive risk values; that is, they provide no information about risks at higher or lower exposure levels.

When low-dose linear extrapolation for cancer effects is supported, particularly for chemicals with direct mutagenic activity or those for which the data indicate a linear component

[19]For most health outcomes (e.g., this would not apply to outcomes related to developmental toxicity), dose-response assessments will be preferably based on studies of chronic exposure. However, analyses will also be conducted for shorter durations, particularly when the evidence base for a PFAS indicates potential risks associated with shorter exposures to the chemical (U.S. EPA, 2002).

below the point of departure (POD), an oral slope factor (OSF) and/or an inhalation unit risk (IUR) will be used to estimate human cancer risks. In general, this will also be the case when no data are available to inform the evaluation of linearity. An OSF is a plausible upper bound lifetime cancer risk from chronic ingestion of a chemical per unit of mass consumed per unit body weight per day (mg/kg-day). An IUR is a plausible upper bound lifetime cancer risk from chronic inhalation of a chemical per unit of air concentration (expressed as ppm or $\mu g/m^3$). In contrast with reference values (RfVs), an OSF or IUR can be used in conjunction with exposure information to predict cancer risk at a given dose.

As discussed in Section 2 "Scoping and Problem Formulation Summary" for these PFAS assessments, Integrated Risk Information System (IRIS) will conduct the assessments with a goal of developing any toxicity values that are reasonably supported by the available data, based on judgments of the evidence drawn during hazard identification and the suitability of studies for dose-response analysis.

The derivation of reference values and cancer risk estimates will depend on the nature of the health hazard conclusions drawn during evidence integration (see Section 10.2). Specifically, EPA generally conducts dose-response assessments and derives cancer values for chemicals that are classified as *carcinogenic* or *likely to be carcinogenic* to humans. When there is *suggestive evidence* of carcinogenicity to humans, EPA generally would not conduct a dose-response assessment or derive a cancer value except when the evidence includes a well-conducted study and quantitative analyses may be useful for some purposes, for example, providing a sense of the magnitude and uncertainty of potential risks, ranking potential hazards, or setting research priorities (U.S. EPA, 2005a). A parallel approach will be taken for potential noncancer health effects in these assessments. Specifically, for noncancer outcomes these assessments will generally include dose-response assessments when the evidence integration judgments indicate there is "sufficient evidence for hazard," with preference given to health effects with stronger evidence scenarios within that category (Section 10.2), and quantitative analyses generally will not be attempted for "insufficient evidence."

11.1. SELECTING STUDIES FOR DOSE-RESPONSE ASSESSMENT

The dose-response assessment will begin with a review of the important health effects highlighted during hazard identification, particularly among the studies of highest quality and that exemplify the study attributes summarized in Table 23. This review will also consider whether there are opportunities for quantitative evidence integration, although it is considered unlikely that the data available to do so will be available for these assessments based on the preliminary literature inventory. Examples of quantitative integration, from simplest to more complex, include (1) the combination of results for an outcome across sex (within a study); (2) characterizing overall toxicity, as in combining effects that comprise a syndrome, or occur on a continuum (e.g., precursors and overt toxicity, benign tumors that progress to malignant tumors); and

(3) meta-analysis or meta-regression of all studies addressing a category of important health effects.

Some studies that were used qualitatively for hazard identification may or may not be considered useful quantitatively for dose-response analysis in these five assessments because of factors like the lack of quantitative measures of exposure or of variability measures for response data. If the needed information cannot be located (e.g., by contacting study authors and making any information publicly available), a semiquantitative analysis (e.g., via no-observed-adverse-effect level [NOAEL]/lowest-observed-adverse-effect level [LOAEL]) will be considered. Studies of low sensitivity may be considered less useful if they failed to detect an effect or reported points of departure with wide confidence limits, but such studies will still be considered for inclusion in a meta-analysis.

Among the studies that support the evidence integration conclusions, those that are most useful for dose-response analysis will generally have at least one exposure level in the region of the dose-response curve near the benchmark response (the response level to be used for deriving toxicity values) to minimize low-dose extrapolation. Such studies will also have more exposure levels and larger sample sizes overall (U.S. EPA, 2012). These attributes support a more complete characterization of the shape of the exposure-response curve and decrease the uncertainty in the associated exposure-response metric (e.g., IUR or RfC) by reducing statistical uncertainty in the POD and minimizing the need for low-dose extrapolation. In addition to these more general considerations, specific issues that may be considered for their potential to impact the feasibility of dose-response modeling for individual data sets are described in more detail in the *Benchmark Dose Technical Guidance* (U.S. EPA, 2012).

Table 21. Attributes used to evaluate studies for deriving toxicity values

Study attributes		Considerations	
		Human studies	**Animal studies**
Rationale for choice of species		Human data are preferred over animal data to **eliminate interspecies extrapolation uncertainties** (e.g., in toxicodynamics, relevance of specific health outcomes to humans).	Animal studies provide supporting evidence when adequate human studies are available and are considered principal studies when adequate human studies are not available. For some hazards, studies of animal species known to respond similarly to humans would be preferred over studies of other species.
Relevance of exposure paradigm	Exposure route	Studies involving **human environmental exposures** (oral, inhalation).	Studies by a route of administration relevant to human environmental exposure are preferred. A validated toxicokinetic model can also be used to extrapolate across exposure routes.
	Exposure durations	When developing a chronic toxicity value, chronic- or subchronic-duration studies are preferred over studies of acute exposure. Exceptions exist, such as when a susceptible population or lifestage is more sensitive in a certain time window (e.g., developmental exposure).	
	Exposure levels	Exposures near the range of typical environmental human exposures are preferred. Studies with a broad exposure range and multiple exposure levels are preferred to the extent that they can provide information about the shape of the exposure-response relationship [see the EPA *Benchmark Dose Technical Guidance*; (U.S. EPA, 2012)] and facilitate extrapolation to more relevant (generally lower) exposures.	
Subject selection		Studies that provide risk estimates in the most susceptible groups are preferred.	
Controls for possible confounding[a]		Studies with a design (e.g., matching procedures, blocking) or analysis (e.g., covariates or other procedures for statistical adjustment) that adequately address the relevant sources of potential critical confounding for a given outcome are preferred.	
Measurement of exposure		Studies that can reliably distinguish between levels of exposure in a time window considered most relevant for a causal effect with respect to the development of the outcome are preferred. Exposure assessment methods that reduce measurement error and methods that provide measurement of exposure at the level of the individual are preferred. Measurements of exposure should not be influenced by knowledge of health outcome status.	Studies providing actual measurements of exposure (e.g., analytical inhalation concentrations versus target concentrations) are preferred. Relevant internal dose measures may facilitate extrapolation to humans, as would availability of a suitable animal PBPK model in conjunction with an animal study reported in terms of administered exposure.

Study attributes	Considerations	
	Human studies	**Animal studies**
Measurement of health outcome(s)	Studies that can reliably distinguish the presence or absence (or degree of severity) of the outcome are preferred. Outcome ascertainment methods using generally accepted, standardized approaches) are preferred.	
	Studies with individual data are preferred in general. Examples include characterizing experimental variability more realistically and characterizing overall incidence of individuals affected by related outcomes (e.g., phthalate syndrome).	
Study size and design	Preference is given to studies using designs reasonably expected to have power to detect responses of suitable magnitude.[b] This does not mean that studies with substantial responses but low power would be ignored, but that they should be interpreted in the context of a confidence interval or variance for the response. Studies that address changes in the number at risk (through decreased survival, loss to follow-up) are preferred.	

[a]An exposure or other variable that is associated with both exposure and outcome but is not an intermediary between the two.
[b]Power is an attribute of the design and population parameters, based on a concept of repeatedly sampling a population; it cannot be inferred post hoc using data from one experiment (Hoenig and Heisey, 2001).

1

11.2. CONDUCTING DOSE-RESPONSE ASSESSMENTS

Consistent with EPA practice, these PFAS assessments will apply a two-step approach for dose-response assessment that distinguishes analysis of the dose-response data in the range of observation from any inferences about responses at lower environmentally relevant exposure levels (U.S. EPA, 2012, 2005a):

1) Within the observed dose range, the preferred approach will be to use dose-response modeling to incorporate as much of the data set as possible into the analysis. This modeling to derive a POD should include an exposure level ideally near the lower end of the range of observation, without significant extrapolation to lower exposure levels (see Section 11.2.1 for more details).

2) As derivation of cancer risk estimates and reference values nearly always involves extrapolation to exposures lower than the POD; the approaches to be applied in these assessments are described in more detail in Section 11.2.2 and Section 11.2.3, respectively.

When sufficient and appropriate human and laboratory animal data are available for the same outcome, human data will be generally preferred for the dose-response assessment because use of human data eliminates the need to perform interspecies extrapolations.

For reference values, these assessments will typically derive a candidate value from each suitable data set, whether in humans or animals (see Section 11.1). Evaluation of these candidate values grouped within a given organ/system will yield a single organ/system-specific value for each organ/system under consideration. Next, evaluation of these organ/system-specific values will result in the selection of a single overall reference value to cover all health outcomes across all organs/systems. While this overall reference value represents the focus of these dose-response assessments, the organ/system-specific values can be useful for subsequent cumulative risk assessments that consider the combined effect of multiple PFAS (or other agents) acting at a common organ/system.

For cancer, if there are multiple tumor sites that can be quantified individually, the final cancer risk estimate(s) will typically address overall cancer risk, to the extent the data allow.

For both cancer and noncancer toxicity values, uncertainties in these estimates will be transparently characterized and discussed.

11.2.1. Dose-Response Analysis in the Range of Observation

Toxicodynamic ("biologically based") modeling is generally preferred when there are sufficient, reliable data to ascertain the MOA and quantitatively support model parameters that represent rates and other quantities associated with the key precursor events of the MOA. Such data, however, do not appear to be available for these five PFAS.

Because a toxicodynamic model will not be available for dose-response assessment, empirical modeling will be used to fit the data (on the apical outcome or a key precursor event) in the range of observation. For this purpose, EPA has developed a standard set of models (http://www.epa.gov/ncea/bmds) that can be applied to typical data sets, including those that are nonlinear. In situations where there are alternative models with significant biological support, the decision maker will be informed by the presentation of these alternatives in the assessment(s) along with the models' strengths and uncertainties. EPA has developed guidance on modeling dose-response data, assessing model fit, selecting suitable models, and reporting modeling results [see the EPA *Benchmark Dose Technical Guidance*; (U.S. EPA, 2012)]. Additional judgment or alternative analyses will be used if the procedure fails to yield reliable results; for example, if the fit is poor, modeling may be restricted to the lower doses, especially if there is competing toxicity at higher doses.

For each modeled response, a POD from the observed data will be estimated to mark the beginning of extrapolation to lower doses. The POD is an estimated dose (expressed in human-equivalent terms) near the lower end of the observed range without significant extrapolation to lower doses. The POD will be used as the starting point for subsequent extrapolations and analyses. For linear extrapolation of cancer risk, the POD will be used to calculate an OSF or IUR, and for nonlinear extrapolation the POD will be used in the calculation of an RfD or RfC.

The response level at which the POD is calculated will be guided by the severity of the endpoint. If linear extrapolation is used, standard values near the low end of the observable range will generally be used (for example, 10% extra risk for cancer bioassay data, 1% for epidemiologic data, lower for rare cancers). For nonlinear approaches, both statistical and biological significance will be considered. For dichotomous data, a response level of 10% extra risk will generally be used for minimally adverse effects, 5% or lower for more severe effects. For continuous data, a response level ideally will be based on an established definition of biologic significance. In the absence of such definition, one control standard deviation from the control mean will generally be used for minimally adverse effects, and one-half standard deviation for more severe effects. The point of departure will be the 95% lower bound on the dose associated with the selected response level.

EPA has developed standard approaches to determine the relevant dose for use in dose-response modeling in the absence of appropriate toxicokinetic modeling. These standard approaches can also aide comparison across exposure patterns and species in the absence of a validated pharmacokinetic (PK) model (see below). The general approaches and considerations to be used to extrapolate PFAS dosimetry from (1) shorter to longer durations within studies, (2) from animals to humans, and (3) across routes of exposure are outlined below:

- Intermittent study exposures will be standardized to a daily average over the duration of exposure. For chronic effects, daily exposures will be averaged over the life span.

Exposures during a critical period, however, will not be averaged over a longer duration (U.S. EPA, 2005a, 1991a). Note that this will typically be done after modeling because the conversion is linear.

- The preferred approach for dosimetry extrapolation from animals to humans will be through PBPK or PK modeling. This approach will be considered first for PFAS and lifestages with existing PBPK models or where an existing model structure can be readily adapted (see Section 6.4 on PBPK modeling).

- Because PK data for the PFAS being evaluated exist in at least one relevant animal species (rats or monkeys) and in humans (see Section 2.4.1), a data-informed extrapolation approach will also be considered for any PFAS that either lacks a PK model or has a model determined to be of inadequate quality. Briefly, the ratio of the elimination half-life in animal to that in the human, $T_{0.5,A}$:$T_{0.5,H}$, or the ratio of the clearance in the human to the animal, CLH:CLA, will be considered for use in converting an oral dose-rate in animals (mg/kg/day) to a human equivalent dose rate (i.e., the human exposure that should yield the same blood concentration as the animal exposure from which it is being extrapolated). Note that clearance and half-life are inversely related. The assessments will consider these metrics as follows:

    - Of these two metrics, $T_{0.5}$ and CL, the half-life is a less complete measure of elimination but one that can be evaluated from more minimal PK data. A half-life can be estimated by observing the decline in an individual's blood concentration of a compound after an exposure has ended. In this way, the total exposure or body burden of the chemical does not have to be known. However, PFAS elimination may go through several phases during which distinct half-lives apply, and the blood concentration that occurs during ongoing exposure may effectively reflect an average among these. The specific approaches and considerations for estimating PFAS half-life are outlined in Section 9.2.1.

    - The clearance, on the other hand, is a measure of average elimination but requires more data to estimate. One must also quantify a companion variable, the volume of distribution (Vd), which in turn requires a measure of total exposure or dose in well-conducted studies. Although more rigorous assessment-specific evaluations will be performed, based on a preliminary review of studies in the literature inventory, the data necessary for the reliable quantification of Vd are expected to be lacking. Specifically, it does not appear that accurate estimates of dose are available in human exposure studies, and the identified animal studies demonstrate considerable interstudy variability in Vd estimates.

- Based on the selection of half-life as the preferred metric and a POD identified from a health-effects study in animals, the human equivalent dose (HED) will be calculated as:

$$HED = (T_{0.5,A[s]}/T_{0.5,H[s]}) \times POD$$

    - Here, the [s] in the subscript indicates that the value may be sex specific. When there are sex-specific values (significant differences between males and females) in both animals and humans, the $T_{0.5}$ values for females would be used to extrapolate health effects in female animals to women, the $T_{0.5}$ values for males used to extrapolate male animal health effects to men. If human data are available to estimate separate half-lives

for women and men, the $T_{0.5}$ for women will likewise be used to estimate HED values in women and the $T_{0.5}$ in men used to estimate HEDs in men. If human data are not sufficient to provide distinct values for men and women, a common $T_{0.5}$ for humans will be used.

- In the absence of PK data/half-lives, oral doses will be scaled allometrically using mg/kg$^{3/4}$ day as the equivalent dose metric across species. Allometric scaling pertains to equivalence across species, not across lifestages, and will not be used to scale doses from adult humans or mature animals to infants or children (U.S. EPA, 2011a, 2005a, 1994). Using this approach, the HED will be calculated as:

$$HED = (BW_H/BW_A)^{0.25} \times POD \text{ (mg/kg-day)}$$

- Inhalation exposures will be scaled using dosimetry models that apply species-specific physiologic and anatomic factors and consider whether the effect occurs at the site of first contact or after systemic circulation (U.S. EPA, 2012, 1994).

- It can be informative to convert doses across exposure routes. If this is done, the assessment will describe the underlying data, algorithms, and assumptions (U.S. EPA, 2005a). Depending on the availability of sufficient data (see Section 9.2) and/or suitable models (see Section 6.4), route-to-route extrapolations in these assessments will be accomplished by using the inhalation exposure rates for PFAS-containing particles predicted using the MPPD model (see Section 9.2) as an ingestion rate in the PK analysis (PBPK/PK model or ADME adjustment), under the assumption that once absorbed into general circulation, the toxic effect is only a function of the body burden or blood concentration.

- In the absence of study-specific data on, for example, intake rates or body weight, the EPA has developed recommended values for use in dose-response analysis (U.S. EPA, 1988).

11.2.2. Extrapolation: Slope Factors and Unit Risk

An OSF or IUR will be used to estimate human cancer risks when low-dose linear extrapolation for cancer effects is supported by the PFAS-specific evidence, particularly for PFAS with direct mutagenic activity or those for which the data indicate a linear component below the POD. Low-dose linear extrapolation will also be used as a default when the data are insufficient to establish the MOA (U.S. EPA, 2005a). If the PFAS-specific data are sufficient to ascertain that one or more modes of action are consistent with low-dose nonlinearity, or to support their biological plausibility, low-dose extrapolation will use the reference-value approach when suitable data are available (U.S. EPA, 2005a); see Section 11.2.3 below.

Differences in susceptibility will be considered for use in deriving multiple slope factors or unit risks, with separate estimates for susceptible populations and lifestages (U.S. EPA, 2005a). If appropriate chemical-specific data on susceptibility from early life exposures are available, then these data will be used to develop cancer slope factors or unit risks that specifically address any potential for differential potency in early lifestages (Farland, 2005; U.S. EPA, 2005a). If such data

are not available, the evidence integration analyses supports a mutagenic MOA for carcinogenicity, and the extrapolation approach is linear, the dose-response assessment will indicate to decision makers that in the development of risk estimates, the default *age-dependent adjustment factors* should be used with the cancer slope factor or unit risk and age-specific estimates of exposure (U.S. EPA, 2005a, b). In this scenario, the final cancer risk value presented in the assessment(s) will reflect this adjustment, with the requisite calculations provided.

The derivation of an OSF and IUR for any of these five PFAS conducted as part of the current assessments will be performed in a manner consistent with EPA guidance.

11.2.3. Extrapolation: Reference Values

Reference value derivation is EPA's most frequently used type of nonlinear extrapolation method, and it will be used in these PFAS assessments for noncancer effects. This approach will also be used for cancer effects if the available data are sufficient to ascertain the MOA and conclude that it is not linear at low doses (see Section 11.2.2). In this case, reference values for each relevant route of exposure will be developed following EPA's established practices (U.S. EPA, 2005a); in general, the reference value will be based not on tumor incidence, but on a key precursor event in the MOA that is necessary for tumor formation. The derivation of an RfD or RfC (if feasible) conducted as part of the assessments for perfluorobutanoic acid (PFBA), PFHxA, perfluorohexanesulfonate (PFHxS), perfluorononanoic acid (PFNA), and perfluorodecanoic acid (PFDA) will be performed in a manner consistent with EPA guidance.

For each data set selected, reference values will be estimated by applying relevant adjustments (i.e., uncertainty factors [UFs]) to the PODs to account for the conditions of the reference value definition. These factors account for human variation, extrapolation from animals to humans, extrapolation to chronic exposure duration, extrapolation to a minimal level of risk (if not observed in the data set), and database deficiencies, as outlined below. Increasingly, data-based adjustments (U.S. EPA, 2014c), probabilistic approaches (Chiu et al., 2018; Chiu and Slob, 2015), and Bayesian methods for characterizing population variability (NAS, 2014) are becoming feasible and may be distinguished from the UF considerations outlined below, if such data exist for these five PFAS. These assessments will discuss the scientific bases (or lack thereof) for each selected UF, including any data-based adjustments based on the following considerations.

- *Animal-to-human extrapolation*: If animal results are used to make inferences about humans, the reference value derivation will incorporate the potential for cross-species differences, which may arise from differences in toxicokinetics or toxicodynamics. The POD will be standardized to equivalent human terms or be based on toxicokinetic or dosimetry modeling, that may range from detailed chemical-specific to default approaches (U.S. EPA, 2014c, 2011a), and a factor of $10^{0.5}$ (rounded to 3) will be applied to account for the remaining uncertainty involving toxicokinetic and toxicodynamic differences. Data-derived adjustments for toxicodynamic differences across species may include qualitative decisions

regarding key science issues (e.g., if, during evaluation of PPARα-dependency, it is concluded that humans are not more sensitive than rodents).

- *Human variation*: The assessments will account for variation in susceptibility across the human population and the possibility that the available data may not represent individuals who are most susceptible to the effect. If appropriate data or models for the effect or for characterizing the internal dose are available, the potential for data-based adjustments for toxicodynamics or toxicokinetics will also be considered (U.S. EPA, 2014c, 2002).[20, 21] When sufficient data are available, an intraspecies UF either less than or greater than 10-fold may be justified (U.S. EPA, 2002). A reduction in this UF will be considered if the POD is derived from or adjusted specifically for susceptible individuals, but not for a general population that includes both susceptible and non-susceptible individuals (U.S. EPA, 2002, 1998, 1996a, 1994, 1991a). In general, when the use of such data or modeling is not supported, a UF with a default value of 10 will be used.

- *LOAEL to NOAEL*: When a POD is based on a LOAEL, the assessment will include an adjustment to an exposure level where such effects are not expected. This can be a matter of great uncertainty if no evidence is available at lower exposures. A factor of 3 or 10 will generally be applied to extrapolate to a lower exposure expected to be without appreciable effects. A factor other than 10 may also be considered, depending on the magnitude and nature of the response and the shape of the dose-response curve (U.S. EPA, 2002, 1998, 1996a, 1994, 1991a).

- *Subchronic-to-chronic exposure*: When using studies of less-than-chronic exposure to make inferences about chronic/lifetime exposure, the assessment will consider whether lifetime exposure could reasonably be interpreted to result in effects at lower levels of exposure, including consideration of the specific health outcome(s) in question. A factor of up to 10 will be considered, depending on the duration of the studies and the nature of the response (U.S. EPA, 2002, 1998, 1994).

- *Database deficiencies*: In addition to the adjustments above, if database deficiencies raise concern that further studies might identify a more sensitive effect, organ system, or lifestage, the assessment will apply a database UF (U.S. EPA, 2002, 1998, 1996a, 1994, 1991a). The size of the factor will depend on the nature of the database deficiency. For example, the EPA typically follows the recommendation that a factor of 10 be applied if both a prenatal toxicity study and a two-generation reproduction study are missing and a factor of $10^{0.5}$ (i.e., 3) if either one or the other is missing (U.S. EPA, 2002). As noted in Section 2.4.5, the evaluation of database completeness for these five PFAS will also consider existing knowledge gained through reviewing other, potentially similar, PFAS to identify data gaps. For example, given the potential for exposure to PFAS to cause developmental effects (based on reviews of perfluorooctanoic acid [PFOA] and perfluorooctane sulfonate

[20]Examples of adjusting the toxicokinetic portion of interhuman variability include the Integrated Risk Information System (IRIS) boron assessment's use of nonchemical-specific kinetic data [e.g., glomerular filtration rate in pregnant humans as a surrogate for boron clearance (U.S. EPA, 2004)] and the IRIS trichloroethylene assessment's use of population variability in trichloroethylene metabolism, via a PBPK model, to estimate the lower 1st percentile of the dose metric distribution for each POD (U.S. EPA, 2011b).

[21]Note that when a PBPK model is available for relating human internal dose to environmental exposure, relevant portions of this UF may be more usefully applied prior to animal-to-human extrapolation, depending on the correspondence of any nonlinearities (e.g., saturation levels) between species.

[PFOS]) and the lack of such studies for PFHxA, consideration of the potential for PFHxA exposure to cause developmental effects might review knowledge gained through the assessment of the other C6 PFAS, PFHxS, or the other short-chain perfluoroalkyl carboxylic acid, perfluorobutanoic acid (PFBA). In such cases, an interpretation of the relatedness between the PFAS of interest and the PFAS used for comparison will inform selection of the factor.

The POD for a particular RfV will be divided by the product of these factors. Based on the RfD/RfC Review (U.S. EPA, 2002) recommendation that any composite factor that exceeds 3,000 represents excessive uncertainty, values with >3,000 UF_C will not be used to derive RfVs. An RfD/RfC may be based on the POD for a single endpoint within a study, or on a collection of related PODs within or across studies, if such biological relationships are substantiated by the evidence.

12. PROTOCOL HISTORY

This section is a placeholder for tracking information on the original protocol release and any potential protocol updates.

REFERENCES

Abbott, BD. (2009). Review of the expression of peroxisome proliferator-activated receptors alpha (PPAR alpha), beta (PPAR beta), and gamma (PPAR gamma) in rodent and human development [Review]. Reprod Toxicol 27: 246-257. http://dx.doi.org/10.1016/j.reprotox.2008.10.001

AK DEC (Alaska Department of Environmental Conservation). (2018). Technical memorandum: action levels for PFAs in water and guidance on sampling groundwater and drinking water. https://dec.alaska.gov/spar/csp/sites/north-pole-refinery/documents/pfas-fact-sheet-1/#axn

ATSDR (Agency for Toxic Substances and Disease Registry). (2018). Toxicological profile for perfluoroalkyls. Draft for public comment. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. https://www.atsdr.cdc.gov/toxprofiles/tp200.pdf

Baduel, C; Paxman, CJ; Mueller, JF. (2015). Perfluoroalkyl substances in a firefighting training ground (FTG), distribution and potential future release. J Hazard Mater 296: 46-53. http://dx.doi.org/10.1016/j.jhazmat.2015.03.007

Barber, JL; Berger, U; Chaemfa, C; Huber, S; Jahnke, A; Temme, C; Jones, KC. (2007). Analysis of per- and polyfluorinated alkyl substances in air samples from Northwest Europe. J Environ Monit 9: 530-541. http://dx.doi.org/10.1039/b701417a

Buck, RC; Franklin, J; Berger, U; Conder, JM; Cousins, IT; de Voogt, P; Jensen, AA; Kannan, K; Mabury, SA; van Leeuwen, SP. (2011). Perfluoroalkyl and polyfluoroalkyl substances in the environment: terminology, classification, and origins [Review]. Integr Environ Assess Manag 7: 513-541. http://dx.doi.org/10.1002/ieam.258

Burri, L; Thoresen, GH; Berge, RK. (2010). The Role of PPARα Activation in Liver and Muscle. PPAR Research 2010. http://dx.doi.org/10.1155/2010/542359

Butenhoff, JL; Bjork, JA; Chang, SC; Ehresman, DJ; Parker, GA; Das, K; Lau, C; Lieder, PH; van Otterdijk, FM; Wallace, KB. (2012a). Toxicological evaluation of ammonium perfluorobutyrate in rats: twenty-eight-day and ninety-day oral gavage studies. Reprod Toxicol 33: 513-530. http://dx.doi.org/10.1016/j.reprotox.2011.08.004

Butenhoff, JL; Chang, SC; Ehresman, DJ; York, RG. (2009). Evaluation of potential reproductive and developmental toxicity of potassium perfluorohexanesulfonate in Sprague Dawley rats. Reprod Toxicol 27: 331-341. http://dx.doi.org/10.1016/j.reprotox.2009.01.004

Butenhoff, JL; Chang, SC; Olsen, GW; Thomford, PJ. (2012b). Chronic dietary toxicity and carcinogenicity study with potassium perfluorooctanesulfonate in Sprague Dawley rats. Toxicology 293: 1-15. http://dx.doi.org/10.1016/j.tox.2012.01.003

Byrne, S; Seguinot-Medina, S; Miller, P; Waghiyi, V; von Hippel, FA; Buck, CL; Carpenter, DO. (2017). Exposure to polybrominated diphenyl ethers and perfluoroalkyl substances in a remote population of Alaska Natives. Environ Pollut 231: 387-395. http://dx.doi.org/10.1016/j.envpol.2017.08.020

CDC (Centers for Disease Control and Prevention). (2009). Fourth national report on human exposure to environmental chemicals. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. http://www.cdc.gov/exposurereport/

CDC (Centers for Disease Control and Prevention). (2015). Fourth national report on human exposure to environmental chemicals, updated tables, February 2015. Atlanta, GA. https://www.cdc.gov/biomonitoring/pdf/FourthReport_UpdatedTables_Feb2015.pdf

CDC (Centers for Disease Control and Prevention). (2018). Fourth national report on human exposure to environmental chemicals, updated tables, March 2018, volume one. Atlanta, GA: U.S. Department of Health and Human Services. https://www.cdc.gov/exposurereport/pdf/FourthReport_UpdatedTables_Volume1_Mar2018.pdf

CDC (Centers for Disease Control and Prevention). (2019a). Fourth National Report on Human Exposure to Environmental Chemicals, Updated Tables, January 2019, Volume 1. https://www.cdc.gov/exposurereport/pdf/FourthReport_UpdatedTables_Volume1_Jan2019-508.pdf

CDC (Centers for Disease Control and Prevention). (2019b). Fourth National Report on Human Exposure to Environmental Chemicals, Updated Tables, January 2019, Volume 2. Atlanta, GA: U.S. Department of Health and Human Services. https://www.cdc.gov/exposurereport/pdf/FourthReport_UpdatedTables_Volume2_Jan2019-508.pdf

Chang, S; Butenhoff, JL; Parker, GA; Coder, PS; Zitzow, JD; Krisko, RM; Bjork, JA; Wallace, KB; Seed, JG. (2018). Reproductive and developmental toxicity of potassium perfluorohexanesulfonate in CD-1 mice. Reprod Toxicol 78: 150-168. http://dx.doi.org/10.1016/j.reprotox.2018.04.007

Chen, WL; Bai, FY; Chang, YC; Chen, PC; Chen, CY. (2018). Concentrations of perfluoroalkyl substances in foods and the dietary exposure among Taiwan general population and pregnant women. J Food Drug Anal 26: 994-1004. http://dx.doi.org/10.1016/j.jfda.2017.12.011

Chiu, WA; Guyton, KZ; Martin, MT; Reif, DM; Rusyn, I. (2018). Use of high-throughput in vitro toxicity screening data in cancer hazard evaluations by IARC Monograph Working Groups. ALTEX 35: 51-64. http://dx.doi.org/10.14573/altex.1703231

Chiu, WA; Slob, W. (2015). A unified probabilistic framework for dose-response assessment of human health effects [Review]. Environ Health Perspect 123: 1241-1254. http://dx.doi.org/10.1289/ehp.1409385

Christensen, KY; Raymond, M; Blackowicz, M; Liu, Y; Thompson, BA; Anderson, HA; Turyk, M. (2017). Perfluoroalkyl substances and fish consumption. Environ Res 154: 145-151. http://dx.doi.org/10.1016/j.envres.2016.12.032

Connecticut DPH (Connecticut Department of Public Health). (2016). Drinking water action level for perfluorinated alkyl substances (PFAS) (pp. 1-9). https://portal.ct.gov/-/media/Departments-and-Agencies/DPH/dph/environmental_health/eoha/Toxicology_Risk_Assessment/DrinkingWaterActionLevelPerfluorinatedAlkylSubstances-PFAS.pdf

Corton, JC; Anderson, SP; Stauber, A. (2000). Central role of peroxisome proliferator-activated receptors in the actions of peroxisome proliferators [Review]. Annu Rev Pharmacol Toxicol 40: 491-518. http://dx.doi.org/10.1146/annurev.pharmtox.40.1.491

Corton, JC; Cunningham, ML; Hummer, BT; Lau, C; Meek, B; Peters, JM; Popp, JA; Rhomberg, L; Seed, J; Klaunig, JE. (2014). Mode of action framework analysis for receptor-mediated toxicity: The peroxisome proliferator-activated receptor alpha (PPARα) as a case study [Review]. Crit Rev Toxicol 44: 1-49. http://dx.doi.org/10.3109/10408444.2013.835784

Corton, JC; Peters, JM; Klaunig, JE. (2018). The PPARα-dependent rodent liver tumor response is not relevant to humans: addressing misconceptions [Review]. Arch Toxicol 92: 83-119. http://dx.doi.org/10.1007/s00204-017-2094-7

Danish EPA (Danish Environmental Protection Agency). (2015). Perfluoroalkylated substances: PFOA, PFOS and PFOSA. Evaluation of health hazards and proposal of a health based quality criterion for drinking water, soil and ground water. (Environmental project No. 1665). https://www2.mst.dk/Udgiv/publications/2015/04/978-87-93283-01-5.pdf

Das, KP; Grey, BE; Rosen, MB; Wood, CR; Tatum-Gibbs, KR; Zehr, RD; Strynar, MJ; Lindstrom, AB; Lau, C. (2015). Developmental toxicity of perfluorononanoic acid in mice. Reprod Toxicol 51: 133-144. http://dx.doi.org/10.1016/j.reprotox.2014.12.012

Dickersin, K. (1990). The existence of publication bias and risk factors for its occurrence. JAMA 263: 1385-1389.

DOE (U.S. Department of Energy). (2016). Temporary emergency exposure limits for chemicals: methods and practice. (DOE-HDBK-1046-2016). https://sp.eota.energy.gov/EM/Shared%20Documents/DOE-HDBK-1046-2016.pdf

DOE (U.S. Department of Energy). (2018). Table 2: Protective Action Criteria (PAC) Rev. 29a based on applicable 60-minute AEGLs, ERPGs, or TEELs. https://sp.eota.energy.gov/pac/docs/Revision_29A_Table2.pdf

Edwards, SW; Tan, YM; Villeneuve, DL; Meek, ME; Mcqueen, CA. (2016). Adverse outcome pathways-organizing toxicological information to improve decision making [Review]. J Pharmacol Exp Ther 356: 170-181. http://dx.doi.org/10.1124/jpet.115.228239

EFSA (European Food Safety Authority). (2008). Perfluorooctane sulfonate (PFOS), perfluorooctanoic acid (PFOA) and their salts. Scientific opinion of the panel on contaminants in the food chain. EFSA J 653: 1-131. http://dx.doi.org/10.2903/j.efsa.2008.653

EFSA (European Food Safety Authority). (2017). Guidance on the use of the weight of evidence approach in scientific assessments. EFSA J 15: 1-69. http://dx.doi.org/10.2903/j.efsa.2017.4971

EIC (ENVIRON International Corporation). (2014). Assessment of POP criteria for specific short-chain perfluorinated alkyl substances. (Project Number: 0134304A). https://fluorocouncil.com/wp-content/uploads/2017/03/ENVIRON-Assessment-of-POP-Criteria-Resources-1.pdf

Fàbrega, F; Kumar, V; Benfenati, E; Schuhmacher, M; Domingo, JL; Nadal, M. (2015). Physiologically based pharmacokinetic modeling of perfluoroalkyl substances in the human body. Toxicol Environ Chem 97: 814-827. http://dx.doi.org/10.1080/02772248.2015.1060976

Farland, WH. (2005). [Memo to Science Policy council regarding implementation of the cancer guidelines and accompanying supplemental guidance - Science Policy Council Cancer Guidelines. Implementation Workgroup communication I: Application of the mode of action framework in mutagenicity determinations for carcinogenicity]. Available online at https://www.epa.gov/sites/production/files/2015-01/documents/cgiwgcommuniation_i.pdf

Filgo, AJ; Quist, EM; Hoenerhoff, MJ; Brix, AE; Kissling, GE; Fenton, SE. (2015). Perfluorooctanoic Acid (PFOA)-induced Liver Lesions in Two Strains of Mice Following Developmental Exposures: PPARα Is Not Required. Toxicol Pathol 43: 558-568. http://dx.doi.org/10.1177/0192623314558463

Foreman, JE; Chang, SC; Ehresman, DJ; Butenhoff, JL; Anderson, CR; Palkar, PS; Kang, BH; Gonzalez, FJ; Peters, JM. (2009). Differential Hepatic Effects of Perfluorobutyrate Mediated by Mouse and Human PPAR-alpha. Toxicol Sci 110: 204-211. http://dx.doi.org/10.1093/toxsci/kfp077

FSANZ (Food Standards Australia New Zealand). (2016). Hazard assessment report: Perfluorooctane sulfonate (PFOS), perfluorooctanoic acid (PFOA), perfluorohexane sulfonate (PFHxS). http://www.health.gov.au/internet/main/publishing.nsf/Content/2200FE086D480353CA2580C900817CDC/$File/6.sd1-Hazard-assessment-report.pdf

Goeden, HM; Greene, CW; Jacobus, JA. (2019). A transgenerational toxicokinetic model and its use in derivation of Minnesota PFOA water guidance. J Expo Sci Environ Epidemiol 29: 183-195. http://dx.doi.org/10.1038/s41370-018-0110-5

Guelfo, JL; Marlow, T; Klein, DM; Savitz, DA; Frickel, S; Crimi, M; Suuberg, EM. (2018). Evaluation and Management Strategies for Per- and Polyfluoroalkyl Substances (PFASs) in Drinking Water Aquifers: Perspectives from Impacted U.S. Northeast Communities. Environ Health Perspect 126: 065001. http://dx.doi.org/10.1289/EHP2727

Guyatt, G; Oxman, AD; Akl, EA; Kunz, R; Vist, G; Brozek, J; Norris, S; Falck-Ytter, Y; Glasziou, P; DeBeer, H; Jaeschke, R; Rind, D; Meerpohl, J; Dahm, P; Schünemann, HJ. (2011). GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 64: 383-394. http://dx.doi.org/10.1016/j.jclinepi.2010.04.026

Guyton, KZ; Chiu, WA; Bateson, TF; Jinot, J; Scott, CS; Brown, RC; Caldwell, JC. (2009). A reexamination of the PPAR-alpha activation mode of action as a basis for assessing human cancer risks of environmental contaminants [Review]. Environ Health Perspect 117: 1664-1672. http://dx.doi.org/10.1289/ehp.0900758

Hall, AP; Elcombe, CR; Foster, JR; Harada, T; Kaufmann, W; Knippel, A; Küttler, K; Malarkey, DE; Maronpot, RR; Nishikawa, A; Nolte, T; Schulte, A; Strauss, V; York, MJ. (2012). Liver hypertrophy: a review of adaptive (adverse and non-adverse) changes--conclusions from the 3rd International ESTP Expert Workshop [Review]. Toxicol Pathol 40: 971-994. http://dx.doi.org/10.1177/0192623312448935

Harbison, RD; Bourgeois, MM; Johnson, GT. (2015). Hamilton & Hardys industrial toxicology (6th ed.). Hoboken, NJ: John Wiley & Sons, Inc. http://www.wiley.com/WileyCDA/WileyTitle/productCd-0470929731,subjectCd-CH50.html

Haug, LS; Thomsen, C; Brantsaeter, AL; Kvalem, HE; Haugen, M; Becher, G; Alexander, J; Meltzer, HM; Knutsen, HK. (2010). Diet and particularly seafood are major sources of perfluorinated compounds in humans. Environ Int 36: 772-778. http://dx.doi.org/10.1016/j.envint.2010.05.016

Health Canada. (2018a). Drinking water screening values: Perfluoroalkylated substances. In Water Talk (Cat. No.: H144-47/2017E-PDF ed.). https://www.canada.ca/en/services/health/publications/healthy-living/water-talk-drinking-water-screening-values-perfluoroalkylated-substances.html

Health Canada. (2018b). Guidelines for Canadian drinking water quality: Guideline technical document - Perfluorooctane sulfonate (PFOS). In Guidelines for Canadian Drinking Water Quality. Ottawa, Ontario: Minister of Health. https://www.canada.ca/en/health-canada/services/publications/healthy-living/guidelines-canadian-drinking-water-quality-guideline-technical-document-perfluorooctane-sulfonate/document.html

Health Canada. (2018c). Guidelines for Canadian drinking water quality: Guideline technical document - Perfluorooctanoic acid (PFOA). In Guidelines for Canadian Drinking Water Quality. Ottawa, Ontario: Minister of Health. https://www.canada.ca/en/health-canada/services/publications/healthy-living/guidelines-canadian-drinking-water-quality-technical-document-perfluorooctanoic-acid/document.html

Heo, JJ; Lee, JW; Kim, SK; Oh, JE. (2014). Foodstuff analyses show that seafood and water are major perfluoroalkyl acids (PFAAs) sources to humans in Korea. J Hazard Mater 279: 402-409. http://dx.doi.org/10.1016/j.jhazmat.2014.07.004

Higgins, JPT; Green, S. (2011). Cochrane handbook for systematic reviews of interventions. Version 5.1.0 (Updated March 2011): The Cochrane Collaboration. Retrieved from http://handbook.cochrane.org/

Hill, AB. (1965). The environment and disease: Association or causation? Proc R Soc Med 58: 295-300.

Hirst, JA; Howick, J; Aronson, JK; Roberts, N; Perera, R; Koshiaris, C; Heneghan, C. (2014). The need for randomization in animal trials: an overview of systematic reviews [Review]. PLoS ONE 9: e98856. http://dx.doi.org/10.1371/journal.pone.0098856

Hoenig, JM; Heisey, DM. (2001). The abuse of power: The pervasive fallacy of power calculations for data analysis. Am Stat 55: 19-24.

Howard, BE; Phillips, J; Miller, K; Tandon, A; Mav, D; Shah, MR; Holmgren, S; Pelch, KE; Walker, V; Rooney, AA; Macleod, M; Shah, RR; Thayer, K. (2016). SWIFT-Review: a text-mining workbench for systematic review. Syst Rev 5: 87. http://dx.doi.org/10.1186/s13643-016-0263-z

Kato, K; Calafat, AM; Needham, LL. (2009). Polyfluoroalkyl chemicals in house dust. Environ Res 109: 518-523. http://dx.doi.org/10.1016/j.envres.2009.01.005

KemI (Swedish Chemicals Agency). (2015). International chemical identification: Nonadecaflourodecanoic acid (PFDA) and its ammonium and sodium salts. https://echa.europa.eu/documents/10162/0dd163d4-56e5-3f3b-8a91-745ec9af2a9f

Kim, SJ; Choi, EJ; Choi, GW; Lee, YB; Cho, HY. (2019). Exploring sex differences in human health risk assessment for PFNA and PFDA using a PBPK model. Arch Toxicol 93: 311-330. http://dx.doi.org/10.1007/s00204-018-2365-y

Kim, SJ; Shin, H; Lee, YB; Cho, HY. (2018). Sex-specific risk assessment of PFHxS using a physiologically based pharmacokinetic model. Arch Toxicol 92: 1113-1131. http://dx.doi.org/10.1007/s00204-017-2116-5

Kim, SK; Kannan, K. (2007). Perfluorinated acids in air, rain, snow, surface runoff, and lakes: relative importance of pathways to contamination of urban lakes. Environ Sci Technol 41: 8328-8334. http://dx.doi.org/10.1021/es072107t

Klaunig, JE; Shinohara, M; Iwai, H; Chengelis, CP; Kirkpatrick, JB; Wang, Z; Bruner, RH. (2015). Evaluation of the chronic toxicity and carcinogenicity of perfluorohexanoic acid (PFHxA) in Sprague-Dawley rats. Toxicol Pathol 43: 209-220. http://dx.doi.org/10.1177/0192623314530532

Krauth, D; Woodruff, TJ; Bero, L. (2013). Instruments for assessing risk of bias and other methodological criteria of published animal studies: a systematic review [Review]. Environ Health Perspect 121: 985-992. http://dx.doi.org/10.1289/ehp.1206389

Laitinen, JA; Koponen, J; Koikkalainen, J; Kiviranta, H. (2014). Firefighters' exposure to perfluoroalkyl acids and 2-butoxyethanol present in firefighting foams. Toxicol Lett 231: 227-232. http://dx.doi.org/10.1016/j.toxlet.2014.09.007

Lau, C. (2015). Perfluorinated compounds: An overview. In JC DeWitt (Ed.), Toxicological effects of perfluoroalkyl and polyfluoroalkyl substances (pp. 1-21). New York: Springer. http://dx.doi.org/10.1007/978-3-319-15518-0_1

Lau, C; Anitole, K; Hodes, C; Lai, D; Pfahles-Hutchens, A; Seed, J. (2007). Perfluoroalkyl acids: a review of monitoring and toxicological findings [Review]. Toxicol Sci 99: 366-394. http://dx.doi.org/10.1093/toxsci/kfm128

Lau, C; Thibodeaux, JR; Hanson, RG; Narotsky, MG; Rogers, JM; Lindstrom, AB; Strynar, MJ. (2006). Effects of perfluorooctanoic acid exposure during pregnancy in the mouse. Toxicol Sci 90: 510-518. http://dx.doi.org/10.1093/toxsci/kfj105

Li, K; Gao, P; Xiang, P; Zhang, X; Cui, X; Ma, LQ. (2017). Molecular mechanisms of PFOA-induced toxicity in animals and humans: Implications for health risks [Review]. Environ Int 99: 43-54. http://dx.doi.org/10.1016/j.envint.2016.11.014

Luebker, DJ; Case, MT; York, RG; Moore, JA; Hansen, KJ; Butenhoff, JL. (2005). Two-generation reproduction and cross-foster studies of perfluorooctanesulfonate (PFOS) in rats. Toxicology 215: 126-148. http://dx.doi.org/10.1016/j.tox.2005.07.018

Mackowiak, B; Hodge, J; Stern, S; Wang, H. (2018). The roles of xenobiotic receptors: Beyond chemical disposition [Review]. Drug Metab Dispos 46: 1361-1371. http://dx.doi.org/10.1124/dmd.118.081042

Macleod, MR. (2013). Systematic reviews of experimental animal studies. Presentation presented at Workshop on weight of evidence; US National Research Council Committee to review the Integrated Risk Information System (IRIS) process, March 27-28, 2013, Washington, DC.

Maloney, EK; Waxman, DJ. (1999). trans-Activation of PPARalpha and PPARgamma by structurally diverse environmental chemicals. Toxicol Appl Pharmacol 161: 209-218. http://dx.doi.org/10.1006/taap.1999.8809

MassDEP (Massachusetts Department of Environmental Protection). (2018). MassDEP ORS Guideline for PFAS: current regulatory limit: per- and polyfluoroalkyl substances (PFAS), including the US EPA UCMR3 analytes. https://www.mass.gov/files/documents/2018/06/11/orsg-pfas-20180608.pdf

MDH (Minnesota Department of Health). (2009). East Metro Perfluorochemical Biomonitoring Pilot Project. (8EHQ-80-373). St. Paul, MN: 3M Company.

MDH (Minnesota Department of Health). (2018). Toxicological summary for: Perfluorobutanoate. Health Based Guidance for Water Health Risk Assessment Unit, Environmental Health Division. https://www.health.state.mn.us/communities/environment/risk/docs/guidance/gw/pfba2summ.pdf

MDH (Minnesota Department of Health). (2019). Health Based Guidance for Water: Toxicological Summary for: Perfluorohexane sulfonate (PFHxS). https://www.health.state.mn.us/communities/environment/risk/docs/guidance/gw/pfhxs.pdf

Meek, ME; Boobis, A; Cote, I; Dellarco, V; Fotakis, G; Munn, S; Seed, J; Vickers, C. (2014). New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis [Review]. J Appl Toxicol 34: 1-18. http://dx.doi.org/10.1002/jat.2949

Mellor, CL; Steinmetz, FP; Cronin, MT. (2016). The identification of nuclear receptors associated with hepatic steatosis to develop and extend adverse outcome pathways [Review]. Crit Rev Toxicol 46: 138-152. http://dx.doi.org/10.3109/10408444.2015.1089471

Moreta, C; Tena, MT. (2014). Determination of perfluorinated alkyl acids in corn, popcorn and popcorn bags before and after cooking by focused ultrasound solid-liquid extraction, liquid chromatography and quadrupole-time of flight mass spectrometry. J Chromatogr A 1355: 211-218. http://dx.doi.org/10.1016/j.chroma.2014.06.018

Morgan, RL; Thayer, KA; Bero, L; Bruce, N; Falck-Ytter, Y; Ghersi, D; Guyatt, G; Hooijmans, C; Langendam, M; Mandrioli, D; Mustafa, RA; Rehfuess, EA; Rooney, AA; Shea, B; Silbergeld, EK; Sutton, P; Wolfe, MS; Woodruff, TJ; Verbeek, JH; Holloway, AC; Santesso, N; Schünemann, HJ. (2016). GRADE: Assessing the quality of evidence in environmental and occupational health. Environ Int 92-93: 611-616. http://dx.doi.org/10.1016/j.envint.2016.01.004

Nabb, DL; Szostek, B; Himmelstein, MW; Mawn, MP; Gargas, ML; Sweeney, LM; Stadler, JC; Buck, RC; Fasano, WJ. (2007). In vitro metabolism of 8-2 fluorotelomer alcohol: interspecies comparisons and metabolic pathway refinement. Toxicol Sci 100: 333-344. http://dx.doi.org/10.1093/toxsci/kfm230

NAS (National Academy of Sciences). (2014). Review of EPA's Integrated Risk Information System (IRIS) process. Washington, DC: National Academies Press (US). http://dx.doi.org/10.17226/18764

New Hampshire DES (New Hampshire Department of Environmental Services). (2019). Summary report on the New Hampshire Department of Environmental Services development of the Maximum Contaminant Levels and Ambient Groundwater Quality Standards for perfluorooctanesulfonic acid (PFOS), perfluorooctanoic acid (PFOA), perfluorononoic acid (PFNA) and perfluorohexanesulfonic acid (PFHxS). (R-WD-19-01). https://www4.des.state.nh.us/nh-pfas-investigation/?p=923

NHMRC (Australian Government National Health and Medical Research Council). (2019). Physical and chemical characteristics: Fact sheets: Per-fluoroalkyl and poly-fluoroalkyl substances (PFAS). In Australian Drinking Water Guidelines.

NICNAS (National Industrial Chemicals Notification and Assessment Scheme). (2014). Public Report: Polyfluorinated polymer in Capstone FS-81 and Capstone TR. (File No. LTD/1406).

NICNAS (National Industrial Chemicals Notification and Assessment Scheme). (2015). Public Report: Efka SL 3239. (File No. LTD/1572).

NIEHS (National Institute of Environmental Health Sciences). (2015). Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration. Office of Health Assessment and Translation (OHAT), National Toxicology Program, U.S Department of Health and Human Services. https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

NJDWQI (New Jersey Drinking Water Quality Institute). (2015). Health-based maximum contaminant level support document: perfluorononanoic acid (PFNA). NJDWQI Health Effects Subcommittee. https://www.state.nj.us/dep/watersupply/pdf/pfna-health-effects.pdf

NJDWQI (New Jersey Drinking Water Quality Institute). (2017). Public review draft: Health-based maximum contaminant level support document: Perfluorooctane sulfonate (PFOS) (CAS #: 1763-23-1; Chemical Formula: C8HF17O3S). NJDWQI Health Effects Subcommittee. https://cswab.org/wp-content/uploads/2018/09/New-Jersey-Scientific-Documentation-Supporting-PFOS-13-ppt.pdf

NLM (National Institutes of Health, National Library of Medicine). (2013). HSDB: Perfluoro-n-nonanoic acid. Available online at https://toxnet.nlm.nih.gov/cgi-bin/sis/search/a?dbs+hsdb:@term+@DOCNO+8040

NLM (National Institutes of Health, National Library of Medicine). (2016). HSDB: Perfluorohexanoic acid. Available online at https://toxnet.nlm.nih.gov/cgi-bin/sis/search/a?dbs+hsdb:@term+@DOCNO+8299

NLM (National Institutes of Health, National Library of Medicine). (2017). HSDB: Perfluorohexanesulfonic acid. Available online at https://toxnet.nlm.nih.gov/cgi-bin/sis/search/a?dbs+hsdb:@term+@DOCNO+8274

Norwegian Environment Agency. (2018). Investigation of sources to PFHxS in the environment. Munich, Germany: BiPRO GmbH. https://www.miljodirektoratet.no/globalassets/publikasjoner/M961/M961.pdf

NRC (National Research Council). (2014). Review of EPA's Integrated Risk Information System (IRIS) process. Washington, DC: The National Academies Press. http://www.nap.edu/catalog.php?record_id=18764

NTP (National Toxicology Program). (2011). NTP Protocol Outline of the Class Study of Perfluorinated Chemicals [perfluorobutane sulfonate (PFBS), perfluorohexane sulfonate potassium salt (PFHSKslt), perfluorooctane sulfonate (PFOS), perfluorohexanoic acid (PFHxA), perfluorooctanoic acid (PFOA), perfluorononanoic acid (PFNA), perfluorodecanoic acid (PFDA) and WY 14643 (WY)] in Harlan Sprague Dawley rats administered daily by gavage for 28-Days [NTP].

NTP (National Toxicology Program). (2015). Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration. U.S. Dept. of Health and Human Services, National Toxicology Program. https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

NTP (National Toxicology Program). (2018a). TOX-96: 1-Perfluorobutanesulfonic acid (375-73-5), potassium perfluorohexanesulfonate (3871-99-6), perfluorooctane sulfonate (1763-23-1), WY-14643 (50892-23-4). Chemical Effects in Biological Systems (CEBS). Available online at

https://manticore.niehs.nih.gov/cebssearch/publication/TOX-96 (accessed August 6, 2018).

NTP (National Toxicology Program). (2018b). TOX-97: Perfluorohexanoic acid (307-24-4), perfluorooctanoic acid (335-67-1), perfluorononanoic acid (375-95-1), perfluorodecanoic acid (335-76-2), WY-14643 (50892-23-4). Chemical Effects in Biological Systems (CEBS). Available online at https://manticore.niehs.nih.gov/cebssearch/publication/TOX-97 (accessed April 24, 2018).

NTP (National Toxicology Program). (2019). TOX-96: Toxicity Report Tables and Curves for Short-term Studies: Perfluorinated Compounds: Sulfonates. Available online at https://tools.niehs.nih.gov/cebs3/views/?action=main.dataReview&bin_id=3874

OECD (Organisation for Economic Co-operation and Development). (2015). Working towards a global emission inventory of PFASS: focus on PFCAS - status quo and the way forward. Paris, France. http://www.oecd.org/chemicalsafety/Working%20Towards%20a%20Global%20Emission%20Inventory%20of%20PFASS.pdf

Papadopoulou, E; Sabaredzovic, A; Namork, E; Nygaard, UC; Granum, B; Haug, LS. (2016). Exposure of Norwegian toddlers to perfluoroalkyl substances (PFAS): The association with breastfeeding and maternal PFAS concentrations. Environ Int 94: 687-694. http://dx.doi.org/10.1016/j.envint.2016.07.006

Peraza, MA; Burdick, AD; Marin, HE; Gonzalez, FJ; Peters, JM. (2006). The toxicology of ligands for peroxisome proliferator-activated receptors (PPAR) [Review]. Toxicol Sci 90: 269-295. http://dx.doi.org/10.1093/toxsci/kfj062

Pérez, F; Llorca, M; Köck-Schulmeyer, M; Škrbić, B; Oliveira, LS; da Boit Martinello, K; Al-Dhabi, NA; Antić, I; Farré, M; Barceló, D. (2014). Assessment of perfluoroalkyl substances in food items at global scale. Environ Res 135: 181-189. http://dx.doi.org/10.1016/j.envres.2014.08.004

Perkins, RG; Butenhoff, JL; Kennedy, GL; Palazzolo, MJ. (2004). 13-week dietary toxicity study of ammonium perfluorooctanoate (APFO) in male rats. Drug Chem Toxicol 27: 361-378. http://dx.doi.org/10.1081/DCT-200039773

Pinney, SM; Biro, FM; Windham, GC; Herrick, RL; Yaghjyan, L; Calafat, AM; Succop, P; Sucharew, H; Ball, KM; Kato, K; Kushi, LH; Bornschein, R. (2014). Serum biomarkers of polyfluoroalkyl compound exposure in young girls in Greater Cincinnati and the San Francisco Bay Area, USA. Environ Pollut 184: 327-334. http://dx.doi.org/10.1016/j.envpol.2013.09.008

Post, GB; Cohn, PD; Cooper, KR. (2012). Perfluorooctanoic acid (PFOA), an emerging drinking water contaminant: a critical review of recent literature [Review]. Environ Res 116: 93-117. http://dx.doi.org/10.1016/j.envres.2012.03.007

Prevedouros, K; Cousins, IT; Buck, RC; Korzeniowski, SH. (2006). Sources, fate and transport of perfluorocarboxylates [Review]. Environ Sci Technol 40: 32-44. http://dx.doi.org/10.1021/es0512475

Rosen, MB; Das, KP; Rooney, J; Abbott, B; Lau, C; Corton, JC. (2017). PPARα-independent transcriptional targets of perfluoroalkyl acids revealed by transcript profiling. Toxicology 387: 95-107. http://dx.doi.org/10.1016/j.tox.2017.05.013

Rotander, A; Kärrman, A; Toms, LM; Kay, M; Mueller, JF; Gómez Ramos, MJ. (2015). Novel fluorinated surfactants tentatively identified in firefighters using liquid chromatography quadrupole time-of-flight tandem mass spectrometry and a case-control approach. Environ Sci Technol 49: 2434-2442. http://dx.doi.org/10.1021/es503653n

Schaider, LA; Balan, SA; Blum, A; Andrews, DQ; Strynar, MJ; Dickinson, ME; Lunderberg, DM; Lang, JR; Peaslee, GF. (2017). Fluorinated compounds in US fast food packaging. Environ Sci Technol Lett 4: 105-111. http://dx.doi.org/10.1021/acs.estlett.6b00435

Schecter, A; Malik-Bass, N; Calafat, AM; Kato, K; Colacino, JA; Gent, TL; Hynan, LS; Harris, TR; Malla, S; Birnbaum, L. (2012). Polyfluoroalkyl compounds in Texas children from birth through 12

years of age. Environ Health Perspect 120: 590-594. http://dx.doi.org/10.1289/ehp.1104325

Schünemann, H; Hill, S; Guyatt, G; Akl, EA; Ahmed, F. (2011). The GRADE approach and Bradford Hill's criteria for causation. J Epidemiol Community Health 65: 392-395. http://dx.doi.org/10.1136/jech.2010.119933

Seacat, AM; Thomford, PJ; Hansen, KJ; Olsen, GW; Case, MT; Butenhoff, JL. (2002). Subchronic toxicity studies on perfluorooctanesulfonate potassium salt in cynomolgus monkeys. Toxicol Sci 68: 249-264. http://dx.doi.org/10.1093/toxsci/68.1.249

Shipley, JM; Hurst, CH; Tanaka, SS; DeRoos, FL; Butenhoff, JL; Seacat, AM; Waxman, DJ. (2004). Trans-activation of PPARalpha and induction of PPARalpha target genes by perfluorooctane-based chemicals. Toxicol Sci 80: 151-160. http://dx.doi.org/10.1093/toxsci/kfh130

Smith, MT; Guyton, KZ; Gibbons, CF; Fritz, JM; Portier, CJ; Rusyn, I; DeMarini, DM; Caldwell, JC; Kavlock, RJ; Lambert, PF; Hecht, SS; Bucher, JR; Stewart, BW; Baan, RA; Cogliano, VJ; Straif, K. (2016). Key characteristics of carcinogens as a basis for organizing data on mechanisms of carcinogenesis [Review]. Environ Health Perspect 124: 713-721. http://dx.doi.org/10.1289/ehp.1509912

SNFA (Swedish National Food Agency). (2018). [Risk management - PFAS in drinking water and fish]. Available online at https://www.livsmedelsverket.se/livsmedel-och-innehall/oonskade-amnen/miljogifter/pfas-poly-och-perfluorerade-alkylsubstanser/riskhantering-pfaa-i-dricksvatten

Soldatow, VY; Lecluyse, EL; Griffith, LG; Rusyn, I. (2013). In vitro models for liver toxicity testing. Toxicology Research 2: 23-39. http://dx.doi.org/10.1039/c2tx20051a

Stahl, LL; Snyder, BD; Olsen, AR; Kincaid, TM; Wathen, JB; McCarty, HB. (2014). Perfluorinated compounds in fish from U.S. urban rivers and the Great Lakes. Sci Total Environ 499: 185-195. http://dx.doi.org/10.1016/j.scitotenv.2014.07.126

Sterne, JAC; Hernán, MA; Reeves, BC; Savović, J; Berkman, ND; Viswanathan, M; Henry, D; Altman, DG; Ansari, MT; Boutron, I; Carpenter, JR; Chan, AW; Churchill, R; Deeks, JJ; Hróbjartsson, A; Kirkham, J; Jüni, P; Loke, YK; Pigott, TD; Ramsay, CR; Regidor, D; Rothstein, HR; Sandhu, L; Santaguida, PL; Schünemann, HJ; Shea, B; Shrier, I; Tugwell, P; Turner, L; Valentine, JC; Waddington, H; Waters, E; Wells, GA; Whiting, PF; Higgins, JPT. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. Br Med J 355: i4919.

Strynar, MJ; Lindstrom, AB. (2008). Perfluorinated compounds in house dust from Ohio and North Carolina, USA. Environ Sci Technol 42: 3751-3756. http://dx.doi.org/10.1021/es7032058

Stump, DG; Holson, JF; Murphy, SR; Farr, CH; Schmit, B; Shinohara, M. (2008). An oral two-generation reproductive toxicity study of S-111-S-WB in rats. Reprod Toxicol 25: 7-20. http://dx.doi.org/10.1016/j.reprotox.2007.10.002

Surma, M; Piskula, M; Wiczkowski, W; Zieliński, H. (2017). The perfluoroalkyl carboxylic acids (PFCAs) and perfluoroalkane sulfonates (PFSAs) contamination level in spices. European Food Research and Technology 243: 297-307. http://dx.doi.org/10.1007/s00217-016-2744-7

Swenberg, JA; Lehman-McKeeman, LD. (1999). Alpha 2-urinary globulin-associated nephropathy as a mechanism of renal tubule cell carcinogenesis in male rats. In CC Capen; E Dybing; JM Rice; JD Wilbourn (Eds.), Species Differences in Thyroid, Kidney, and Urinary Bladder Carcinogenesis (pp. 95-118). Lyon, France: International Agency for Research on Cancer. http://apps.who.int/bookorders/anglais/detart1.jsp?sesslan=1&codlan=1&codcol=73&codcch=147

Takacs, ML; Abbott, BD. (2007). Activation of mouse and human peroxisome proliferator-activated receptors (alpha, beta/delta, gamma) by perfluorooctanoic acid and perfluorooctane sulfonate. Toxicol Sci 95: 108-117. http://dx.doi.org/10.1093/toxsci/kfl135

TCEQ (Texas Commission on Environmental Quality). (2015). TCEQ guidelines to develop toxicity factors. (Revised RG-442). Austin, TX. https://www.tceq.texas.gov/assets/public/comm_exec/pubs/rg/rg-442.pdf

TCEQ (Texas Commission on Environmental Quality). (2013). Tiered Development of Human Health PCLs. https://www.tceq.texas.gov/assets/public/comm_exec/pubs/rg/rg-366-trrp-22.pdf

TCEQ (Texas Commission on Environmental Quality). (2018). TRRP protective concentration levels: April 2018 PCL and supporting tables. Retrieved from https://www.tceq.texas.gov/remediation/trrp/trrppcls.html

Thomford, PJ. (2002). 104-week dietary chronic toxicity and carcinogenicity study with perfluorooctane sulfonic acid potassium salt (PFOS; T-6295) in rats. Study No. 6329-183. Madison, WI: Convance Laboratories.

U.S. EPA (U.S. Environmental Protection Agency). (1988). Recommendations for and documentation of biological values for use in risk assessment [EPA Report] (pp. 1-395). (EPA/600/6-87/008). Cincinnati, OH: U.S. Environmental Protection Agency, Office of Research and Development, Office of Health and Environmental Assessment. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=34855

U.S. EPA (U.S. Environmental Protection Agency). (1991a). Guidelines for developmental toxicity risk assessment (pp. 1-71). (EPA/600/FR-91/001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=23162

U.S. EPA (U.S. Environmental Protection Agency). (1991b). Guidelines for developmental toxicity risk assessment. Fed Reg 56: 63798-63826.

U.S. EPA (U.S. Environmental Protection Agency). (1992). Guidelines for Exposure Assessment [EPA Report]. In Risk Assessment Forum. (EPA/600/Z-92/001). Washington, DC: U. S. Environmental Protection Agency. https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=15263

U.S. EPA (U.S. Environmental Protection Agency). (1994). Methods for derivation of inhalation reference concentrations and application of inhalation dosimetry [EPA Report]. (EPA/600/8-90/066F). Research Triangle Park, NC: U.S. Environmental Protection Agency, Office of Research and Development, Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office. https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=71993&CFID=51174829&CFTOKEN=25006317

U.S. EPA (U.S. Environmental Protection Agency). (1996a). Guidelines for reproductive toxicity risk assessment (pp. 1-143). (EPA/630/R-96/009). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www.epa.gov/sites/production/files/2014-11/documents/guidelines_repro_toxicity.pdf

U.S. EPA (U.S. Environmental Protection Agency). (1996b). Guidelines for reproductive toxicity risk assessment. Fed Reg 61: 56274-56322.

U.S. EPA (U.S. Environmental Protection Agency). (1998). Guidelines for neurotoxicity risk assessment [EPA Report] (pp. 1-89). (EPA/630/R-95/001F). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. http://www.epa.gov/risk/guidelines-neurotoxicity-risk-assessment

U.S. EPA (U.S. Environmental Protection Agency). (2002). A review of the reference dose and reference concentration processes (pp. 1-192). (EPA/630/P-02/002F). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. http://www.epa.gov/osa/review-reference-dose-and-reference-concentration-processes

U.S. EPA (U.S. Environmental Protection Agency). (2004). Toxicological review of boron and compounds. In support of summary information on the Integrated Risk Information System (IRIS) [EPA Report]. (EPA/635/04/052). Washington, DC: U.S. Environmental Protection Agency, IRIS. http://nepis.epa.gov/exe/ZyPURL.cgi?Dockey=P1006CK9.txt

U.S. EPA (U.S. Environmental Protection Agency). (2005a). Guidelines for carcinogen risk assessment [EPA Report] (pp. 1-166). (EPA/630/P-03/001F). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www.epa.gov/sites/production/files/2013-09/documents/cancer_guidelines_final_3-25-05.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2005b). Supplemental guidance for assessing susceptibility from early-life exposure to carcinogens [EPA Report]. (EPA/630/R-03/003F). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www3.epa.gov/airtoxics/childrens_supplement_final.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2011a). Recommended use of body weight 3/4 as the default method in derivation of the oral reference dose (pp. 1-50). (EPA/100/R11/0001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum, Office of the Science Advisor. https://www.epa.gov/risk/recommended-use-body-weight-34-default-method-derivation-oral-reference-dose

U.S. EPA (U.S. Environmental Protection Agency). (2011b). Toxicological review of trichloroethylene (CASRN 79-01-6) in support of summary information on the Integrated Risk Information System (IRIS) [EPA Report]. (EPA/635/R-09/011F). Washington, DC. https://cfpub.epa.gov/ncea/iris/iris_documents/documents/toxreviews/0106tr.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2012). Benchmark dose technical guidance. (EPA/100/R-12/001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www.epa.gov/risk/benchmark-dose-technical-guidance

U.S. EPA (U.S. Environmental Protection Agency). (2014a). Draft: Health effects document for perfluorooctane sulfonate (PFOS) [EPA Report]. (EPA/822/R-14/002). Washington, DC: U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division. https://peerreview.versar.com/epa/pfoa/pdf/Health-Effects-Document-for-Perfluorooctane-Sulfonate-(PFOS).pdf

U.S. EPA (U.S. Environmental Protection Agency). (2014b). Draft: Health effects document for Perfluorooctanoic Acid (PFOA) [EPA Report]. (EPA Document Number: 822R14001). Office of Water, Health and Ecological Criteria Division. https://peerreview.versar.com/epa/pfoa/pdf/Health-Effects-Document-for-Perfluorooctanoic-Acid-(PFOA).pdf

U.S. EPA (U.S. Environmental Protection Agency). (2014c). Guidance for applying quantitative data to develop data-derived extrapolation factors for interspecies and intraspecies extrapolation [EPA Report]. (EPA/100/R-14/002F). Washington, DC: Risk Assessment Forum, Office of the Science Advisor. https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2014d). Health effects document for perfluorooctane sulfonate (PFOS): Draft [EPA Report]. (EPA/822/R-14/002). Washington, DC: U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division. https://peerreview.versar.com/epa/pfoa/pdf/Health-Effects-Document-for-Perfluorooctane-Sulfonate-(PFOS).pdf

U.S. EPA (U.S. Environmental Protection Agency). (2016a). Drinking water health advisory for perfluorooctane sulfonate (PFOS) [EPA Report]. (EPA 822-R-16-004). Washington, DC: U.S. Environmental Protection Agency, Office of Water. https://www.regulations.gov/document?D=EPA-HQ-OW-2014-0138-0038

U.S. EPA (U.S. Environmental Protection Agency). (2016b). Drinking water health advisory for perfluorooctanoic acid (PFOA). (EPA 822-R-16-005). Washington, DC: U.S. Environmental Protection Agency, Office of Water. https://www.regulations.gov/document?D=EPA-HQ-OW-2014-0138-0041

U.S. EPA (U.S. Environmental Protection Agency). (2016c). Health effects support document for perfluorooctane sulfonate (PFOS) [EPA Report]. (EPA 822-R-16-002). Washington, DC: U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division. https://www.epa.gov/sites/production/files/2016-05/documents/pfos_hesd_final_508.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2016d). Health effects support document for perfluorooctanoic acid (PFOA) [EPA Report]. (EPA 822-R-16-003). Washington, DC: U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division. https://www.epa.gov/sites/production/files/2016-05/documents/pfoa_hesd_final-plain.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2016e). The Third Unregulated Contaminant Monitoring Rule. Available online at https://www.epa.gov/dwucmr/third-unregulated-contaminant-monitoring-rule

U.S. EPA (U.S. Environmental Protection Agency). (2017a). Guidance to assist interested persons in developing and submitting draft risk evaluations under the Toxic Substances Control Act. (EPA/740/R17/001). Washington, DC: U.S Environmental Protection Agency, Office of Chemical Safety and Pollution Prevention. https://www.epa.gov/sites/production/files/2017-06/documents/tsca_ra_guidance_final.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2017b). The Third Unregulated Contaminant Monitoring Rule (UCMR 3): Data summary, January 2017. (EPA 815-S-17-001). Washington, DC: U.S Environmental Protection Agency, Office of Water. https://www.epa.gov/sites/production/files/2017-02/documents/ucmr3-data-summary-january-2017.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2018a). Chemistry Dashboard. Washington, DC. Retrieved from https://comptox.epa.gov/dashboard

U.S. EPA (U.S. Environmental Protection Agency). (2018b). An umbrella Quality Assurance Project Plan (QAPP) for PBPK models [EPA Report]. (ORD QAPP ID No: B-0030740-QP-1-1). Research Triangle Park, NC.

U.S. EPA (U.S. Environmental Protection Agency). (2019). ChemView [Database]. Retrieved from https://chemview.epa.gov/chemview

van Otterdijk, FM. (2007). Repeated dose 28-day oral toxicity study with MTDID-8391 by daily gavage in the rat, followed by a 21-day recovery period. (Study Number 06-226). Maplewood, MN: 3M.

Verner, MA; Ngueta, G; Jensen, ET; Fromme, H; Vöelkel, W; Nygaard, UC; Granum, B; Longnecker, MP. (2016). A simple pharmacokinetic model of prenatal and postnataleExposure to Perfluoroalkyl Substances (PFASs). Environ Sci Technol 50: 978-986. http://dx.doi.org/10.1021/acs.est.5b04399

Vesterinen, HM; Sena, ES; Egan, KJ; Hirst, TC; Churolov, L; Currie, GL; Antonic, A; Howells, DW; Macleod, MR. (2014). Meta-analysis of data from animal studies: a practical guide. J Neurosci Methods 221: 92-102. http://dx.doi.org/10.1016/j.jneumeth.2013.09.010

Viberg, H; Eriksson, P. (2017). Chapter 43 - Perfluorooctane sulfonate and perfluorooctanoic acid. In RC Gupta (Ed.), Reproductive and Developmental Toxicology (2nd ed., pp. 811-827). San Diego, CA: Academic Press. http://dx.doi.org/10.1016/B978-0-12-804239-7.00043-3

von der Trenck, KT; Konietzka, R; Biegel-Engler, A; Brodsky, J; Hädicke, A; Quadflieg, A; Stockerl, R; Stahl, T. (2018). Significance thresholds for the assessment of contaminated groundwater:

perfluorinated and polyfluorinated chemicals [Review]. Environ Sci Eur 30: 19. http://dx.doi.org/10.1186/s12302-018-0142-4

VT DOH (Vermont Department of Health). (2018). [Memorandum to Emily Boedecker from Mark A. Levine regarding drinking water health advisory for five PFAs (per- and polyfluorinated alkyl substances)]. Available online at http://www.healthvermont.gov/sites/default/files/documents/pdf/ENV_DW_PFAS_HealthAdvisory.pdf

Wambaugh, JF; Setzer, RW; Pitruzzello, AM; Liu, J; Reif, DM; Kleinstreuer, NC; Wang, NC; Sipes, N; Martin, M; Das, K; Dewitt, JC; Strynar, M; Judson, R; Houck, KA; Lau, C. (2013). Dosimetric anchoring of in vivo and in vitro studies for perfluorooctanoate and perfluorooctanesulfonate. Toxicol Sci 136: 308-327. http://dx.doi.org/10.1093/toxsci/kft204

Wang, Y; Wang, L; Li, J; Liang, Y; Ji, H; Zhang, J; Zhou, Q; Jiang, G. (2014). The mechanism of immunosuppression by perfluorooctanoic acid in BALB/c mice. Toxicology Research 3: 205-213. http://dx.doi.org/10.1039/c3tx50096a

Wilhelm, M; Bergmann, S; Dieter, HH. (2010). Occurrence of perfluorinated compounds (PFCs) in drinking water of North Rhine-Westphalia, Germany and new approach to assess drinking water contamination by shorter-chained C4-C7 PFCs. Int J Hyg Environ Health 213: 224-232. http://dx.doi.org/10.1016/j.ijheh.2010.05.004

Wolf, CJ; Rider, CV; Lau, C; Abbott, BD. (2014). Evaluating the additivity of perfluoroalkyl acids in binary combinations on peroxisome proliferator-activated receptor-α activation. Toxicology 316: 43-54. http://dx.doi.org/10.1016/j.tox.2013.12.002

Wolf, CJ; Takacs, ML; Schmid, JE; Lau, C; Abbott, BD. (2008). Activation of mouse and human peroxisome proliferator-activated receptor alpha by perfluoroalkyl acids of different functional groups and chain lengths. Toxicol Sci 106: 162-171. http://dx.doi.org/10.1093/toxsci/kfn166

WS (Weston Solutions Inc). (2007). Remedial investigation report. Phase 2. Fluorochemical (FC) data assessment report for the Cottage Grove, MN site. St. Paul, MN: 3M Corporate Toxicology. https://www.pca.state.mn.us/sites/default/files/pfc-cottagegrove-remedialinvestigationreport.pdf

Yoo, H; Washington, JW; Jenkins, TM; Ellington, JJ. (2011). Quantitative determination of perfluorochemicals and fluorotelomer alcohols in plants from biosolid-amended fields using LC/MS/MS and GC/MS. Environ Sci Technol 45: 7985-7990. http://dx.doi.org/10.1021/es102972m

Zhao, P; Xia, X; Dong, J; Xia, N; Jiang, X; Li, Y; Zhu, Y. (2016). Short- and long-chain perfluoroalkyl substances in the water, suspended particulate matter, and surface sediment of a turbid river. Sci Total Environ 568: 57-65. http://dx.doi.org/10.1016/j.scitotenv.2016.05.221

APPENDIX A. SUMMARY OF EXISTING TOXICITY VALUE INFORMATION FOR PERFLUOROBUTANOIC ACID (PFBA), PERFLUOROHEXANOIC ACID (PFHXA), PERFLUORONONANOIC ACID (PFNA), AND PERFLUORODECANOIC ACID (PFDA)

Table A-1. Details on derivation of the available health effect reference values for inhalation exposure to selected per- and polyfluoroalkyl substances (PFAS)

	Reference value name	Duration	PFAS	Reference value		Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
				(mg/m^3)	(ppm)							
Emergency response	**PAC-3**	1 h	PFBA	3.3×10^{1}	3.6×10^{0}	NR	NR	NR		NR	PAC values derived via an approach developed by the Department of Energy (DOE, 2016)	Final (DOE, 2018)
	PAC-2	1 h	PFBA	5.5×10^{0}	6.0×10^{-1}	Based on PAC-3	--	--	--	--	Based on PAC-3[a]	
	PAC-1	1 h	PFBA	5.0×10^{-1}	5.5×10^{-2}	Based on PAC-2	--	--	--	--	Based on PAC-2[b]	
General	**TCEQ RfC**	Chronic	PFBA	1.0×10^{-2}	1.1×10^{-3}	NR	NR	NR		NR	RfCs developed with TCEQ's protocol (TCEQ, 2012)	Final (TCEQ, 2018)
			PFDA	5.3×10^{-5}	2.5×10^{-6}	NR	NR	NR		NR		
			PFNA	2.8×10^{-5}	1.4×10^{-6}	NR	NR	NR		NR		

	Reference value name	Duration	PFAS	Reference value		Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
				(mg/m^3)	(ppm)							
			PFHxS	1.3×10^{-5}	7.8×10^{-7}	NR	NR	NR		NR		

[a]PAC-2 = PAC-3 ÷ 6 = 33 mg/m^3 ÷ 6 = 5.5 mg/m^3.

[b]PAC-1 = PAC-1 ÷ 11 = 5.5 mg/m^3 ÷ 11 = 0.5 mg/m^3.

NR = not reported; PAC = Protective Action Criteria; PFAS = per- and polyfluoroalkyl substances; PFBA = perfluorobutanoic acid; PFDA = perfluorodecanoic acid; PFHxS = perfluorohexanesulfonate; PFNA = perfluorononanoic acid; RfC = inhalation reference concentration; TCEQ = Texas Commission on Environmental Quality; UF_A = animal to human variability; UF_D = database uncertainty; UF_H = interhuman variability; UF_L = LOAEL-to-NOAEL adjustment; UF_S = subchronic-to-chronic adjustment.

Table A-2. Details on derivation of the available health effect reference values for oral exposure to selected per- and polyfluoroalkyl substances (PFAS)

Reference value name	Duration	PFAS	Reference value (mg/kg-d)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
EPA RfD (OW)	Chronic	PFOA	2×10^{-5}	Decreased ossification and accelerated male puberty in F1 mice	1 mg/kg-d, 38 mg/L serum 0.0053 mg/kg-d	LOAEL $LOAEL_{HED}$	Lau et al. (2006)	Total UF = 300 $UF_A = 3$ $UF_H = 10$ $UF_L = 10$ $UF_S = 1$ $UF_D = 1$	Average serum concentration derived using a PBPK model developed by Wambaugh et al. (2013) HED adjusted[a]	Final (U.S. EPA, 2016b)
		PFOS	2×10^{-5}	Reduced body wt. in F2 rats	0.1 mg/kg-d, 6.26 µg/mL 0.00051 mg/kg-d	NOAEL $NOAEL_{HED}$	Luebker et al. (2005)	Total UF = 30 $UF_A = 3$ $UF_H = 10$ $UF_L = 1$ $UF_S = 1$ $UF_D = 1$	Average serum concentration derived using a PBPK model developed by Wambaugh et al. (2013) HED adjusted[b]	Final (U.S. EPA, 2016a)
EFSA TDI	Chronic	PFOS	1.5×10^{-4}	Changes in lipids and thyroid hormones in monkeys	0.03 mg/kg-d	NOAEL	Seacat et al. (2002)	Total UF = 200 $UF_A = 10$ $UF_H = 10$ $UF_D = 2$		Final (EFSA, 2008)

Reference value name	Duration	PFAS	Reference value (mg/kg-d)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
NH DES RfD	Chronic	PFHxS	9.3×10^{-6}	Reduced litter size in mice exposed for 14 d	0.3 mg/kg-d, 27,200 ng/L serum 90.7 ng/mL	NOAEL Target human serum level	Chang et al. (2018)	Total UF = 300 $UF_A = 3$ $UF_H = 10$ MF[c] = 10	Target human serum level = BMDL ÷ UF Calculated[d]	Final (New Hampshire DES, 2019)
		PFNA	2.5×10^{-6}	Increased relative liver wts. in mice	4,900 ng/L serum 16.3 ng/mL serum	BMDL Target human serum level	Das et al. (2015)	Total UF = 300 $UF_A = 3$ $UF_H = 10$ MF[e] = 10	Target human serum level = BMDL ÷ UF Calculated[f]	
TCEQ RfD	Chronic	PFBA	2.9×10^{-3}	NR	NR	NR		NR	RfDs developed with TCEQ's protocol (TCEQ, 2012)	Final (TCEQ, 2018)
		PFDA	1.5×10^{-5}	NR	NR	NR		NR		
		PFHxA	3.8×10^{-6}	NR	NR	NR		NR		
		PFHxS	3.8×10^{-6}	NR	NR	NR		NR		
		PFNA	1.2×10^{-5}	NR	NR	NR		NR		
Australia Dept. of Health TDI	Chronic	Combined PFOS and PFHxS	2×10^{-5}	Decreased body wt. gain in F0 female rats	0.1 mg/kg-d, 7.14 µg/mL 0.0006 mg/kg-d	NOAEL $NOAEL_{HED}$	Luebker et al. (2005)	Total UF = 30 $UF_A = 3$ $UF_H = 10$	HED Adjusted[g]	Final (FSANZ, 2016)

Reference value name	Duration	PFAS	Reference value (mg/kg-d)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
Danish EPA TDI	Chronic	Combined PFOS, PFBS, PFDS, PFHpA, PFHxA, PFHxS, and PFNA	3×10^{-5}	Liver lesions in male rats	0.033 mg/kg-d 0.0008 mg/kg-d	$BMDL_{10}$ $BMDL_{HED}$	Thomford (2002)	Total UF = 30 $UF_A = 3$ $UF_H = 10$	HED Adjusted[h] Pharmaco-kinetic adjustments based on those in U.S. EPA (2014a)	Final (Danish EPA, 2015)

[a]$LOAEL_{HED}$ = LOAEL × volume of distribution × (ln2 ÷ $t_{1/2}$) = 38 mg/L × 0.17 L/kg × (0.693 ÷ 839.5 days) = 0.0053 mg/kg-day.
[b]$NOAEL_{HED}$ = NOAEL × volume of distribution × (ln2 ÷ $t_{1/2}$) = 6.26 µg/mL × 0.23 L/kg × (0.693 ÷ 1,971 days) = 0.00051 mg/kg-day.
[c]The modifying factor of 10 is to account for "the limited number of studies on PFHxS, both animal and epidemiological, as well as uncertainty for associated effects on other physiological processes including the thyroid system."
[d]RfD = THSL × volume of distribution × (ln2 ÷ $t_{1/2}$) = 90.7 ng/mL × 0.287 L/kg × (0.693 ÷ [5.3 years × 365 days/year]) × 1,000 mL/L = 9.3 ng/kg-day.
[e]The modifying factor of 10 is to account for "the limited number of studies on PFNA, specifically the lack of information for a serum half-life in humans, as well as uncertainty for associated effects on other physiological processes including the immune system."
[f]RfD = THSL × volume of distribution × (ln2 ÷ $t_{1/2}$) = 16.3 ng/mL × 0.2 L/kg × (0.693 ÷ [2.5 years × 365 days/year]) × 1,000 mL/L = 2.5 ng/kg-day.
[g]TDI = NOAEL × volume of distribution × (ln2 ÷ $t_{1/2}$) = 7.14 µg/mL × 0.23 L/kg × (0.693 ÷ 1,971 days) = 0.0006 mg/kg-day.
[h]$BMDL_{HED}$ = $BMDL_{10}$ ÷ ([volume of distribution × (ln2 ÷ $t_{1/2\ Rat}$)] ÷ [volume of distribution × (ln2 ÷ $t_{1/2\ Human}$)]) = 0.033 mg/kg-day ÷ ([0.23 L/kg × (0.693 ÷ 48 days)] ÷ [0.23 L/kg × (0.693 ÷ 1,971 days)]) = 0.0008 mg/kg-day.
BMDL = benchmark dose lower confidence limit; EFSA = European Food Safety Authority; EPA = Environmental Protection Agency; HED = human equivalent dose; LOAEL = lowest-observed-adverse-effect level; MDH = Minnesota Department of Health; MF = modifying factor; NH DES = New Hampshire Department of Environmental Services; NOAEL = no-observed-adverse-effect level; NR = not reported; OW = Office of Water; PBPK = physiologically based pharmacokinetic; PFAS = per- and polyfluoroalkyl substances; PFBA = perfluorobutanoic acid; PFBS = perfluorobutane sulfonate; PFDA = perfluorodecanoic acid; PFHxA = perfluorohexanoic acid; PFHxS = perfluorohexanesulfonate; PFNA = Perfluorononanoic acid; PFOS = perfluorooctane sulfonate; PK = pharmacokinetic; RfD = oral reference dose; SD = standard deviation; TCEQ = Texas Commission on Environmental Quality; TDI = tolerable daily intake; THSL = target human serum level; UF_A = animal to human variability; UF_D = database uncertainty; UF_H = interhuman variability; UF_L = LOAEL-to-NOAEL adjustment; UF_S = subchronic-to-chronic adjustment.

Table A-3. Details on derivation of the available drinking water standards for selected per- and polyfluoroalkyl substances (PFAS)

Reference value name	Duration	PFAS	Reference value (µg/L)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
EPA DWEL (OW)	Chronic	PFOA	0.37	Based on RfD	--	--	--	--	Based on RfD[a]	Final (U.S. EPA, 2016b)
		PFOS	0.37							Final (U.S. EPA, 2016a)
EPA HA (OW)	Chronic	Combined PFOA and PFOS	0.07	Based on PFOA and PFOS DWELs	--	--	--	--	Based on PFOA and PFOS DWELs[b]	Final (U.S. EPA, 2016a, b)
MDH HRL[c]	1–30 d Subchronic Chronic	PFBA	7 7 7	Decreased cholesterol, serum total thyroxine, and dialysis free thyroxine and increased relative thyroid wt. in rats	3.01 mg/kg-day 0.38 mg/kg-day 0.0038 mg/kg-day	$BMDL_{1sd}$ $BMDL_{HED}$ RfD	(van Otterdijk, 2007)	Total UF = 100 UF_A = 3 UF_H = 10 UF_D = 3	HED Adjusted[d] Calculated[e]	Final (MDH, 2018)

Reference value name	Duration	PFAS	Reference value (µg/L)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
MDH HBV	1–30 d Subchronic Chronic	PFHxS	0.047 0.047 0.047	Decreased free and total T4 and triiodothyronine (T3), changes in cholesterol levels, and increased hepatic focal necrosis in rats	32.4 mg/L serum 0.00292 mg/kg-day 0.0000097 mg/kg-day	BMDL $BMDL_{HED}$ RfD	(NTP, 2019)	Total UF = 300 UF_A = 3 UF_H = 10 UF_D = 10	HED Adjusted[f] Developed from RfD using toxicokinetic model (Goeden et al., 2019)	Final (MDH, 2019)
Mass. ORSG	Chronic	Combined PFOA, PFOS, PFNA, PFHpA, and PFHxS	0.07	Adopted EPA HA for PFOA and PFOS	--	--	--	--	Adopted EPA HA for PFOA and PFOS	Final (MassDEP, 2018)
CT DPH Drinking Water Action Level	Chronic	Combined PFOA, PFOS, PFNA, PFHpA, and PFHxS	0.07	Adopted EPA HA for PFOA and PFOS	--	--	--	--	Adopted EPA HA for PFOA and PFOS	Final (Connecticut DPH, 2016)
AK DEC Action Level	Chronic	Combined PFOA, PFOS, PFNA, PFHpA, and PFHxS	0.07	Adopted EPA HA for PFOA and PFOS	--	--	--	--	Adopted EPA HA for PFOA and PFOS	Final (AK DEC, 2018)
NH DES DWEL	Chronic	PFHxS	0.1691	Based on RfD	--	--	--	--	Based on RfD[g]	Final (New Hampshire DES, 2019)
		PFNA	0.0455	Based on RfD	--	--	--	--	Based on RfD[h]	
NH DES MCL	Chronic	PFHxS	0.085	Based on DWEL	--	--	--	--	Based on DWEL[i]	

Reference value name	Duration	PFAS	Reference value (µg/L)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
		PFNA	0.023	Based on DWEL	--	--	--	--	Based on DWEL[j]	
NJ DEP MCL	Chronic	PFNA	0.013	Increased liver wts. in pregnant mice	4,900 ng/mL serum 4,900 ng/L serum	BMDL Target human serum level	Das et al. (2015)	Total UF = 1,000 UF_A = 3 UF_H = 10 UF_S = 10 UF_D = 3	Target human serum level = BMDL ÷ UF Calculated[k]	Final (NJDWQI, 2015)
VT DEC Drinking Water HA	Chronic	Combined PFOA, PFOS, PFHxS, PFHpA, PFNA	0.02	Based on EPA OW RfD	--	--	--	--	Based on EPA OW RfD[l]	Final (VT DOH, 2018)
TCEQ Tier 1 PCL	Chronic	PFBA	71	NR	NR	NR		NR	PCL derived in accordance with TCEQ protocol (TCEQ, 2013)	Final (TCEQ, 2018)
		PFDA	0.37	NR	NR	NR		NR		
		PFHxA	0.093	NR	NR	NR		NR		
		PFHxS	0.093	NR	NR	NR		NR		
		PFNA	0.29	NR	NR	NR		NR		
Australia Dept. of Health Drinking Water Guideline	Chronic	Combined PFOS and PFHxS	0.07	Based on TDI	--	--	--	--	Based on TDI[m]	Final (NHMRC, 2019)

Reference value name	Duration	PFAS	Reference value (µg/L)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
Danish EPA QC	Chronic	Combined PFOA, PFOS, PFBS, PFDS, PFHpA, PFHxA, PFHxS, and PFNA	0.1	Based on TDI	--	--	--	--	Based on TDI[n]	Final (Danish EPA, 2015)
GFS	Chronic	PFBA	10	Hepatocellular hypertrophy and thyroid hyperplasia and hypertrophy in rats exposed for 90 d	6 mg/kg-d 3 µg/kg-d	NOAEL $NOAEL_{HED}$	Butenhoff et al. (2012a) and van Otterdijk (2007)	Total UF = 250 UF_A = 2.5 UF_H = 10 UF_S = 10	HED adjusted[o] Calculated[p]	Final (von der Trenck et al., 2018)
		PFNA	0.06	Liver lesions and hepatocellular hypertrophy in rats	0.025 mg/kg-d 0.0167 µg/kg-d	LOAEL $LOAEL_{HED}$	Stump et al. (2008)	Total UF = 30 UF_L = 3 UF_D = 10	HED adjusted[q] Calculated[r]	
		PFHxA	6	Renal toxicity and lower urine pH in rats exposed for 104 wk	15 mg/kg-d 1.84 µg/kg-d	NOAEL $NOAEL_{HED}$	Klaunig et al. (2015), NICNAS (2015), EIC (2014), and NICNAS (2014)	Total UF = 25 UF_A = 2.5 UF_H = 10	HED adjusted[s] Calculated[t]	

Reference value name	Duration	PFAS	Reference value (µg/L)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
		PFHxS	0.1	Increased liver wt., hepatocellular hypertrophy, thyroid hyperplasia, and decreased hematocrit in rats exposed for 42 d	1 mg/kg-d 0.03 µg/kg-d	NOAEL $NOAEL_{HED}$	Butenhoff et al. (2009)	Total UF = 375 UF_A = 2.5 UF_H = 10 UF_S = 15	HED adjusted[u] Calculated[v]	
UBA HRIV[w]	Chronic	PFDA	0.1	NR	NR	NR		NR	Based on estimated half-lives	Provisional (von der Trenck et al., 2018)
		PFHxA	1	NR	NR	NR		NR		Provisional (Wilhelm et al., 2010)
		PFHxS	0.3	NR	NR	NR		NR		
Health Canada MAC	Chronic	PFOA	0.2	Hepatocellular hypertrophy and increased liver wt. in rats	0.05 mg/kg-day 0.000521 mg/kg-day 0.000021 mg/kg-day	$BMDL_{10}$ $BMDL_{HED}$ TDI	(Perkins et al., 2004)	Total UF = 25 UF_A = 2.5 UF_H = 10	HED Adjusted[x] Calculated[y]	Final (Health Canada, 2018c)

Reference value name	Duration	PFAS	Reference value (µg/L)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status
		PFOS	0.6	Hepatocellular hypertrophy in rats	0.021 mg/kg-day 0.0015 mg/kg-day 0.00006 mg/kg-day	NOAEL $NOAEL_{HED}$ TDI	(Butenhoff et al., 2012b)	Total UF = 25 UF_A = 2.5 UF_H = 10	HED Adjusted[z] Calculated[aa]	Final (Health Canada, 2018b)
Health Canada Drinking Water Screening Value	Chronic	PFBA	30	NR	NR	NR		NR	NR	Final (Health Canada, 2018a)
		PFNA	0.02	NR	NR	NR		NR	NR	
		PFHxA	0.2	Adopted MAC for PFOA	--	--	--	--	Adopted MAC for PFOA	
		PFHxS	0.6	Adopted MAC for PFOS	--	--	--	--	Adopted MAC for PFOS	
Sweden MTL	Chronic	Combined PFOA, PFBS, PFHpA, PFHxA, PFHxS, PFHpA, and PFPeA	0.09	Based on EFSA TDI for PFOS	--	--	--	--	Based on EFSA TDI for PFOS[bb]	Final (SNFA, 2018)[cc]

[a]DWEL = RfD ÷ water intake rate = 0.00002 mg/kg-day ÷ 0.054 L/kg-day = 0.00037 mg/L.
[b]HA = DWEL × RSC = 0.00037 mg/L × 0.2 = 0.00007 mg/L.
[c]The HRL was adopted as a drinking water quality standard in Italy.
[d]$BMDL_{HED}$ = $BMDL_{1sd}$ ÷ ($t_{1/2\ Human}$ ÷ $t_{1/2\ Male\ Rat}$) = 3.01 mg/kg-day ÷ 72 h ÷ 9.22 h = 0.38 mg/kg-day.
[e]HRL = RfD × RSC ÷ short-term intake rate = 0.0038 mg/kg-day × 0.5 ÷ 0.285 L/kg-day = 0.0067 mg/L × 1,000 µg/mg = 7 µg/L.
[f]$BMDL_{HED}$ = BMDL × volume of distribution × (ln2 ÷ $t_{1/2}$) = 32.4 mg/L × 0.25 L/kg × 0.693 ÷ 1935 days = 0.00292 mg/kg-day.
[g]DWEL = RfD ÷ water intake rate = 9.3 ng/kg-day ÷ 0.055 L/kg-day = 169.1 ng/L.
[h]DWEL = RfD ÷ water intake rate = 2.5 ng/kg-day ÷ 0.055 L/kg-day = 45.5 ng/L.
[i]MCL = DWEL × RSC = 169.1 ng/L × 0.5 = 85 ng/L.
[j]MCL = DWEL × RSC = 45.5 ng/L × 0.5 = 23 ng/L.
[k]MCL = THSL × RSC ÷ serum:drinking water ratio = 4,900 ng/L serum × 0.5 ÷ (200 ng/L serum ÷ 1 ng/L drinking water) = 13 ng/L.
[l]DWHA = RFD × RSC × (1 ÷ BW_AIR) = 0.00002 mg/kg-day × 0.2 × 1 ÷ 0.175 L/kg-day × 1,000 µg/mL = 0.07 µg/L.
[m]DW guideline = TDI × BW × RSC ÷ water intake rate = 0.00002 mg/kg-day × 70 kg × 0.1 ÷ 2 L/day = 0.00007 mg/L × 1,000 µg/mL = 0.07 µg/L.
[n]QC = TDI × = RSC ÷ water intake rate = 0.03 µg/kg-day × 0.1 ÷ 0.03 L/kg-day = 0.1 µg/L.
[o]$NOAEL_{HED}$ = NOAEL ÷ ($t_{1/2\ Human}$ ÷ $t_{1/2\ Rat}$) ÷ UF = 6 mg/kg-day ÷ 8 ÷ 250 = 0.003 mg/kg-day.
[p]GFS = $NOAEL_{HED}$ × BW × RSC ÷ water intake rate = 3 µg/kg-day × 70 kg × 0.1 ÷ 2 L/day = 10 µg/L.
[q]$NOAEL_{HED}$ = NOAEL ÷ ($t_{1/2\ Human}$ ÷ $t_{1/2\ Rat}$) ÷ UF = 0.025 mg/kg-day ÷ 50 ÷ 30 = 0.0000167 mg/kg-day.
[r]GFS = $NOAEL_{HED}$ × BW × RSC ÷ water intake rate = 0.0167 µg/kg-day × 70 kg × 0.1 ÷ 2 L/day = 0.058 µg/L.
[s]$NOAEL_{HED}$ = NOAEL ÷ ($t_{1/2\ Human}$ ÷ $t_{1/2\ Rat}$) ÷ UF = 15 mg/kg-day ÷ 327 ÷ 25 = 0.00184 mg/kg-day.
[t]GFS = $NOAEL_{HED}$ × BW × RSC ÷ water intake rate = 1.84 µg/kg-day × 70 kg × 0.1 ÷ 2 L/day = 6.4 µg/L.
[u]$NOAEL_{HED}$ = NOAEL ÷ ($t_{1/2\ Human}$ ÷ $t_{1/2\ Rat}$) ÷ UF = 1 mg/kg-day ÷ 90 ÷ 375 = 0.00003 mg/kg-day.
[v]GFS = $NOAEL_{HED}$ × BW × RSC ÷ water intake rate = 0.03 µg/kg-day × 70 kg × 0.1 ÷ 2 L/day = 0.1037 µg/L.
[w]The HRIV for PFHxA was adopted as a drinking water quality standard in Italy.
[x]$BMDL_{HED}$ = $BMDL_{10}$ ÷ species- and dose-specific adjustment factor = 0.05 mg/kg-day ÷ 96 = 0.000521 mg/kg-day.
[y]MAC = TDI × BW × RSC ÷ water intake rate = 0.000021 mg/kg-day × 70 kg × 0.2 ÷ 1.5 L/day = 0.0002 mg/L.
[z]$NOAEL_{HED}$ = NOAEL ÷ species- and dose-specific adjustment factor = 0.021 mg/kg-day ÷ 14 = 0.0015 mg/kg-day.
[aa]MAC = TDI × BW × RSC ÷ water intake rate = 0.00006 mg/kg-day × 70 kg × 0.2 ÷ 1.5 L/day = 0.0006 mg/L.
[bb]The Danish EPA's summary of other existing regulations states that Sweden's MTL was derived "considering an exposure scenario where 10% of [the PFOS TDI] was allocated to the consumption of infant formula based on drinking water."
[cc]Source document not available in English.
AK DEC = Alaska Department of Environmental Conservation; BMDL = benchmark dose lower confidence limit; CT DPH = Connecticut Department of Health; DWEL = drinking water equivalent level; EFSA = European Food Safety Authority; EPA = Environmental Protection Agency; GFS = *Geringfügigkeitsschwelle*, a significance threshold for the assessment of contaminated groundwater; HA = health advisory; HBV = health-based value; HED = human equivalent dose; HRIV = health risk indication value; HRL = health risk limit; LOAEL = lowest-observed-adverse effect- level; MAC = maximum acceptable concentration; Mass. ORSG = MassDEP Office of Research and Standards Guidelines; MCL = maximum contaminant level; MDH = Minnesota Department of Health; MTL = maximal tolerable level; NH DES = New Hampshire Department of Environmental Services; NJ DEP = New Jersey Department of Environmental Protection; NOAEL = no-observed-adverse effect- level; NR = not reported; OW = Office of Water; PBPK = physiologically based pharmacokinetic; PCL = protective

Reference value name	Duration	PFAS	Reference value (µg/L)	Health effect	Point of departure	Qualifier	Source	Uncertainty factors	Notes on derivation	Review status

concentration level; PFAS = per and polyfluoroalkyl substances; PFBA = perfluorobutanoic acid; PFBS = perfluorobutane sulfonate; PFDA = perfluorodecanoic acid; PFHxA = perfluorohexanoic acid; PFHxS = perfluorohexanesulfonate; PFNA = Perfluorononanoic acid; PFOS = perfluorooctane- sulfonate; PK = pharmacokinetic; QC = quality criterion RfD = oral reference dose; RSC = relative source contribution; TCEQ = Texas Commission on Environmental Quality; TDI = tolerable daily intake; THSL = target human serum level; UBA = *Umweltbundesamt*, the German Federal Environment Agency; UF_A = animal to human variability; UF_D = database uncertainty; UF_H = interhuman variability; UF_L = LOAEL-to-NOAEL adjustment; UF_S = subchronic-to-chronic adjustment; VT DEC = Vermont Department of Environmental Conservation.

1

APPENDIX B. SEARCH AND SCREENING STRATEGIES

Table B-1. Perfluorobutanoic acid (PFBA) database search strategy

Search	Search strategy	Dates of search
PubMed		
Search terms	375-22-4[rn] OR "Heptafluoro-1-butanoic acid"[tw] OR "Heptafluorobutanoic acid"[tw] OR "Heptafluorobutyric acid"[tw] OR "Kyselina heptafluormaselna"[tw] OR "Perfluorobutanoic acid"[tw] OR "Perfluorobutyric acid"[tw] OR "Perfluoropropanecarboxylic acid"[tw] OR "2,2,3,3,4,4,4-heptafluoro-Butanoic acid"[tw] OR "Butanoic acid, 2,2,3,3,4,4,4-heptafluoro-"[tw] OR "Butanoic acid, heptafluoro-"[tw] OR "Perfluoro-n-butanoic acid"[tw] OR "Perfluorobutanoate"[tw] OR "2,2,3,3,4,4,4-Heptafluorobutanoic acid"[tw] OR "Butyric acid, heptafluoro-"[tw] OR "Fluorad FC 23"[tw] OR "H 0024"[tw] OR "NSC 820"[tw] OR ((PFBA[tw] OR "FC 23"[tw] OR HFBA[tw]) AND (fluorocarbon*[tw] OR fluorotelomer*[tw] OR polyfluoro*[tw] OR perfluoro-*[tw] OR perfluoroa*[tw] OR perfluorob*[tw] OR perfluoroc*[tw] OR perfluorod*[tw] OR perfluoroe*[tw] OR perfluoroh*[tw] OR perfluoron*[tw] OR perfluoroo*[tw] OR perfluorop*[tw] OR perfluoros*[tw] OR perfluorou*[tw] OR perfluorinated[tw] OR fluorinated[tw] OR PFAS[tw] OR PFOS[tw] OR PFOA[tw]))	No date limit–7/19/2017
Literature update search terms	(((375-22-4[rn] OR "Heptafluoro-1-butanoic acid"[tw] OR "Heptafluorobutanoic acid"[tw] OR "Heptafluorobutyric acid"[tw] OR "Kyselina heptafluormaselna"[tw] OR "Perfluorobutanoic acid"[tw] OR "Perfluorobutyric acid"[tw] OR "Perfluoropropanecarboxylic acid"[tw] OR "2,2,3,3,4,4,4-heptafluoro-Butanoic acid"[tw] OR "Butanoic acid, 2,2,3,3,4,4,4-heptafluoro-"[tw] OR "Butanoic acid, heptafluoro-"[tw] OR "Perfluoro-n-butanoic acid"[tw] OR "Perfluorobutanoate"[tw] OR "2,2,3,3,4,4,4-Heptafluorobutanoic acid"[tw] OR "Butyric acid, heptafluoro-"[tw] OR "Fluorad FC 23"[tw] OR "H 0024"[tw] OR "NSC 820"[tw] OR ((PFBA[tw] OR "FC 23"[tw] OR HFBA[tw]) AND (fluorocarbon*[tw] OR fluorotelomer*[tw] OR polyfluoro*[tw] OR perfluoro-*[tw] OR perfluoroa*[tw] OR perfluorob*[tw] OR perfluoroc*[tw] OR perfluorod*[tw] OR perfluoroe*[tw] OR perfluoroh*[tw] OR perfluoron*[tw] OR perfluoroo*[tw] OR perfluorop*[tw] OR perfluoros*[tw] OR perfluorou*[tw] OR perfluorinated[tw] OR fluorinated[tw] OR PFAS[tw] OR PFOS[tw] OR PFOA[tw])) AND ("2017/08/01"[PDAT] : "2018/02/14"[PDAT])	8/1/2017–2/14/2018

Search	Search strategy	Dates of search
Web of Science		
Search terms	TS="Heptafluoro-1-butanoic acid" OR TS="Heptafluorobutanoic acid" OR TS="Heptafluorobutyric acid" OR TS="Kyselina heptafluormaselna" OR TS="Perfluorobutanoic acid" OR TS="Perfluorobutyric acid" OR TS="Perfluoropropanecarboxylic acid" OR TS="2,2,3,3,4,4,4-heptafluoro-Butanoic acid" OR TS="Butanoic acid, 2,2,3,3,4,4,4-heptafluoro-" OR TS="Butanoic acid, heptafluoro-" OR TS="Perfluoro-n-butanoic acid" OR TS="Perfluorobutanoate" OR TS="2,2,3,3,4,4,4-Heptafluorobutanoic acid" OR TS="Butyric acid, heptafluoro-" OR TS="Fluorad FC 23" OR TS="H 0024" OR TS="NSC 820" OR (TS=(PFBA OR "FC 23" OR HFBA) AND TS=(fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro-* OR perfluoroa* OR perfluorob* OR perfluoroc* OR perfluorod* OR perfluoroe* OR perfluoroh* OR perfluoron* OR perfluoroo* OR perfluorop* OR perfluoros* OR perfluorou* OR perfluorinated OR fluorinated OR PFAS OR PFOS OR PFOA))	No date limit–7/20/2017
Literature update search terms	((TS="Heptafluoro-1-butanoic acid" OR TS="Heptafluorobutanoic acid" OR TS="Heptafluorobutyric acid" OR TS="Kyselina heptafluormaselna" OR TS="Perfluorobutanoic acid" OR TS="Perfluorobutyric acid" OR TS="Perfluoropropanecarboxylic acid" OR TS="2,2,3,3,4,4,4-heptafluoro-Butanoic acid" OR TS="Butanoic acid, 2,2,3,3,4,4,4-heptafluoro-" OR TS="Butanoic acid, heptafluoro-" OR TS="Perfluoro-n-butanoic acid" OR TS="Perfluorobutanoate" OR TS="2,2,3,3,4,4,4-Heptafluorobutanoic acid" OR TS="Butyric acid, heptafluoro-" OR TS="Fluorad FC 23" OR TS="H 0024" OR TS="NSC 820") OR TS=(PFBA OR "FC 23" OR HFBA) AND TS=(fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro-* OR perfluoroa* OR perfluorob* OR perfluoroc* OR perfluorod* OR perfluoroe* OR perfluoroh* OR perfluoron* OR perfluoroo* OR perfluorop* OR perfluoros* OR perfluorou* OR perfluorinated OR fluorinated OR PFAS OR PFOS OR PFOA)) AND PY=2017-2018	2017–2018
Toxline		
Search terms	(375-22-4 [rn] OR "heptafluoro-1-butanoic acid" OR "heptafluorobutanoic acid" OR "heptafluorobutyric acid" OR "kyselina heptafluormaselna" OR "perfluorobutanoic acid" OR "perfluorobutyric acid" OR "perfluoropropanecarboxylic acid" OR "2,2,3,3,4,4,4-heptafluoro-butanoic acid" OR "butanoic acid 2 2 3 3 4 4 4-heptafluoro-" OR "butanoic acid heptafluoro-" OR "perfluoro-n-butanoic acid" OR "perfluorobutanoate" OR "2,2,3,3,4,4,4-heptafluorobutanoic acid" OR "butyric acid heptafluoro-" OR "fluorad fc 23" OR "h 0024" OR "nsc 820" OR ((pfba OR "fc 23" OR hfba) AND (fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro* OR perfluorinated OR fluorinated OR pfas OR pfos OR pfoa))) AND (ANEUPL [org] OR BIOSIS [org] OR CIS [org] OR DART [org] OR EMIC [org] OR EPIDEM [org] OR HEEP [org] OR HMTC [org] OR IPA [org] OR RISKLINE [org] OR MTGABS [org] OR NIOSH [org] OR NTIS [org] OR PESTAB [org] OR PPBIB [org]) AND NOT PubMed [org] AND NOT pubdart [org]	No date limit–7/20/2017

Search	Search strategy	Dates of search
Literature update search terms	@AND+@OR+("heptafluoro-1-butanoic acid"+"heptafluorobutanoic+acid"+"heptafluorobutyric+acid"+"kyselina+heptafluormaselna"+"perfluorobutanoic+acid"+"perfluorobutyric+acid"+"perfluoropropanecarboxylic +acid"+"2 2 3 3 4 4 4-heptafluoro-butanoic+acid"+"butanoic+acid+2 2 3 3 4 4 4-heptafluoro-"+"butanoic+acid+heptafluoro-"+"perfluoro-n-butanoic acid"+"perfluorobutanoate"+"2 2 3 3 4 4 4-heptafluorobutanoic+acid"+"butyric+acid+heptafluoro-"+"fluorad+fc+23"+"h0024"+"nsc+820"+@TERM+@rn+375-22-4("pfba"+"fc+23"+"hfba"))+(fluorocarbon*+ fluorotelomer*+polyfluoro*+perfluoro*+perfluorinated+fluorinated+pfas+pfos+pfoa)+@RANGE+yr+2017+2018	2017–2018
TSCATS		
Search terms	375-22-4[rn] AND tscats[org]	No date limit–7/20/2017

Table B-2. Perfluorodecanoic acid (PFDA) database search strategy

Search	Search strategy	Dates of search
PubMed		
Search terms	335-76-2[rn] OR "Ndfda"[tw] OR "Nonadecafluoro-n-decanoic acid"[tw] OR "Nonadecafluorodecanoic acid"[tw] OR "Perfluoro-n-decanoic acid"[tw] OR "Perfluorodecanoic acid"[tw] OR "2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-Decanoic acid"[tw] OR "Decanoic acid, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-"[tw] OR "Decanoic acid, nonadecafluoro-"[tw] OR "Perfluorodecanoate"[tw] OR "PFDeA"[tw] OR "PFDcA"[tw] OR ("PFDA"[tw] AND (fluorocarbon*[tw] OR fluorotelomer*[tw] OR polyfluoro*[tw] OR perfluoro-*[tw] OR perfluoroa*[tw] OR perfluorob*[tw] OR perfluoroc*[tw] OR perfluorod*[tw] OR perfluoroe*[tw] OR perfluoroh*[tw] OR perfluoron*[tw] OR perfluoroo*[tw] OR perfluorop*[tw] OR perfluoros*[tw] OR perfluorou*[tw] OR perfluorinated[tw] OR fluorinated[tw] OR PFAS[tw] OR PFOS[tw] OR PFOA[tw]))	No date limit–7/26/2017
Literature update search terms	((335-76-2[rn] OR "Ndfda"[tw] OR "Nonadecafluoro-n-decanoic acid"[tw] OR "Nonadecafluorodecanoic acid"[tw] OR "Perfluoro-n-decanoic acid"[tw] OR "Perfluorodecanoic acid"[tw] OR "2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-Decanoic acid"[tw] OR "Decanoic acid, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-"[tw] OR "Decanoic acid, nonadecafluoro-"[tw] OR "Perfluorodecanoate"[tw] OR "PFDeA"[tw] OR "PFDcA"[tw] OR ("PFDA"[tw] AND (fluorocarbon*[tw] OR fluorotelomer*[tw] OR polyfluoro*[tw] OR perfluoro-*[tw] OR perfluoroa*[tw] OR perfluorob*[tw] OR perfluoroc*[tw] OR perfluorod*[tw] OR perfluoroe*[tw] OR perfluoroh*[tw] OR perfluoron*[tw] OR perfluoroo*[tw] OR perfluorop*[tw] OR perfluoros*[tw] OR perfluorou*[tw] OR perfluorinated[tw] OR fluorinated[tw] OR PFAS[tw] OR PFOS[tw] OR PFOA[tw])) AND ("2017/08/01"[Date - Publication] : "2018/03/01"[Date - Publication])	8/1/2017–2/14/2018

Search	Search strategy	Dates of search
Web of Science		
Search terms	TS="PFDeA" OR TS="PFDcA" OR TS="Ndfda" OR TS="Nonadecafluoro-n-decanoic acid" OR TS="Nonadecafluorodecanoic acid" OR TS="Perfluoro-n-decanoic acid" OR TS="Perfluorodecanoic acid" OR TS="2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-Decanoic acid" OR TS="Decanoic acid, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-" OR TS="Decanoic acid, nonadecafluoro-" OR TS="Perfluorodecanoate" OR (TS=PFDA AND TS=(fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro-* OR perfluoroa* OR perfluorob* OR perfluoroc* OR perfluorod* OR perfluoroe* OR perfluoroh* OR perfluoron* OR perfluoroo* OR perfluorop* OR perfluoros* OR perfluorou* OR perfluorinated OR fluorinated)) OR (TS=PFDA AND TS=(fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro-* OR perfluoroa* OR perfluorob* OR perfluoroc* OR perfluorod* OR perfluoroe* OR perfluoroh* OR perfluoron* OR perfluoroo* OR perfluorop* OR perfluoros* OR perfluorou* OR perfluorinated OR fluorinated OR PFAS OR PFOS OR PFOA	No date limit–7/26/2017
Literature update search terms	TS="PFDeA" OR TS="PFDcA" OR TS="Ndfda" OR TS="Nonadecafluoro-n-decanoic acid" OR TS="Nonadecafluorodecanoic acid" OR TS="Perfluoro-n-decanoic acid" OR TS="Perfluorodecanoic acid" OR TS="2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-Decanoic acid" OR TS="Decanoic acid, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-" OR TS="Decanoic acid, nonadecafluoro-" OR TS="Perfluorodecanoate" OR (TS=PFDA AND TS=(fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro-* OR perfluoroa* OR perfluorob* OR perfluoroc* OR perfluorod* OR perfluoroe* OR perfluoroh* OR perfluoron* OR perfluoroo* OR perfluorop* OR perfluoros* OR perfluorou* OR perfluorinated OR fluorinated)) OR (TS=PFDA AND TS=(fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro-* OR perfluoroa* OR perfluorob* OR perfluoroc* OR perfluorod* OR perfluoroe* OR perfluoroh* OR perfluoron* OR perfluoroo* OR perfluorop* OR perfluoros* OR perfluorou* OR perfluorinated OR fluorinated OR PFAS OR PFOS OR PFOA)) AND PY=2017-2018	2017–2018
Toxline		
Search terms	(335-76-2 [rn] OR "2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluorodecanoic acid" OR "2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-decanoic acid" OR "decanoic acid 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-nonadecafluoro-" OR "decanoic acid nonadecafluoro-" OR "nonadecafluoro-n-decanoic acid" OR "nonadecafluorodecanoic acid" OR "perfluoro-1-nonanecarboxylic acid" OR "perfluoro-n-decanoic acid" OR "perfluorocapric acid" OR "perfluorodecanoate" OR "perfluorodecanoic acid" OR "ndfda" OR "PFDeA" OR "PFDcA" OR (pfda AND (fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro* OR perfluorinated OR fluorinated OR pfas OR pfos OR pfoa))) AND (ANEUPL [org] OR BIOSIS [org] OR CIS [org] OR DART [org] OR EMIC [org] OR EPIDEM [org] OR HEEP [org] OR HMTC [org] OR IPA [org] OR RISKLINE [org] OR MTGABS [org] OR NIOSH [org] OR NTIS [org] OR PESTAB [org] OR PPBIB [org]) AND NOT PubMed [org] AND NOT pubdart [org]	No date limit–7/21/2017

Search	Search strategy	Dates of search
Literature update search terms		2017−2018
TSCATS		
Search terms	335-76-2[rn] AND TSCATS[org]	No date limit−7/21/2017

1

Table B-3. Perfluorononanoic acid (PFNA) database search strategy

Search	Search strategy	Dates of search
PubMed		
Search terms	"375-95-1"[rn] OR "2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluorononanoic acid"[tw] OR "Nonanoic acid, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluoro-"[tw] OR "Nonanoic acid, heptadecafluoro-"[tw] OR "Perfluoro-n-nonanoic acid"[tw] OR "Perfluorononan-1-oic acid"[tw] OR "Perfluorononanoate"[tw] OR "Perfluorononanoic acid"[tw] OR "Perfluorononanonic acid"[tw] OR "Perfluoropelargonic acid"[tw] OR "heptadecafluorononanoic acid"[tw] OR (("PFNA"[tw] OR "C 1800"[tw]) AND (fluorocarbon*[tw] OR fluorotelomer*[tw] OR polyfluoro*[tw] OR perfluoro-*[tw] OR perfluoroa*[tw] OR perfluorob*[tw] OR perfluoroc*[tw] OR perfluorod*[tw] OR perfluoroe*[tw] OR perfluoroh*[tw] OR perfluoron*[tw] OR perfluoroo*[tw] OR perfluorop*[tw] OR perfluoros*[tw] OR perfluorou*[tw] OR perfluorinated[tw] OR fluorinated[tw] OR PFAS[tw] OR PFOS[tw] OR PFOA[tw]))	No date limit–7/26/2017
Literature update and additional PFNA synonyms search terms	((("2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluorononanoic acid" [tw] OR "Nonanoic acid, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluoro-" [tw] OR "Nonanoic acid, heptadecafluoro-" [tw] OR "Perfluoro-n-nonanoic acid" [tw] OR "Perfluorononan-1-oic acid" [tw] OR "Perfluorononanoate" [tw] OR "Perfluorononanoic acid" [tw] OR "Perfluorononanonic acid" [tw] OR "Perfluoropelargonic acid" [tw] OR "heptadecafluorononanoic acid" [tw] OR "PFNA" [tw] OR "C 1800" [tw] OR "Methyl-n1-Perfluorononanoic acid" [tw] OR "PFNA-n1CH3" [tw] OR "EINECS 206-801-3" [tw] OR "Heptadecafluornonansaeure" [tw] OR "Heptadekafluornonansaeure" [tw] OR "Perfluornonansaeure" [tw] OR "Perfluorononanoic acid (PFNA)" [tw] OR "UNII-5830Z6S63M" [tw] OR "perfluoro-n-nonanoic acid" [tw] OR "perfluorononan-1-oic acid" [tw] OR "perfluorononanoic acid" [tw] OR "Ammonium Perfluorononanoate" [tw] OR "Ammonium perfluorononanoate" [tw] OR "PFNA-H3N" [tw])) AND ("2017/01/01"[Date - Publication] : "3000"[Date - Publication])	1/2017–4/2018
Web of Science		
Search terms	((TS=PFNA OR TS="C 1800") AND TS=(fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro-* OR perfluoroa* OR perfluorob* OR perfluoroc* OR perfluorod* OR perfluoroe* OR perfluoroh* OR perfluoron* OR perfluoroo* OR perfluorop* OR perfluoros* OR perfluorou* OR perfluorinated OR fluorinated OR PFAS OR PFOS OR PFOA)) OR TS="2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluorononanoic acid" OR TS="Nonanoic acid, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluoro-" OR TS="Nonanoic acid, heptadecafluoro-" OR TS="Perfluoro-n-nonanoic acid" OR TS="Perfluorononan-1-oic acid" OR TS="Perfluorononanoate" OR TS="Perfluorononanoic acid" OR TS="Perfluorononanonic acid" OR TS="Perfluoropelargonic acid" OR TS="heptadecafluorononanoic acid”	No date limit–7/26/2017

Search	Search strategy	Dates of search
Literature update and additional PFNA synonyms search terms	(TS="PFNA" OR TS="C 1800" OR TS="2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluorononanoic acid" OR TS="Nonanoic acid, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluoro-" OR TS="Nonanoic acid, heptadecafluoro-" OR TS="Methyl-n1-Perfluorononanoic acid" OR TS="PFNA-n1CH3" OR TS="EINECS 206-801-3" OR TS="Heptadecafluornonansaeure" OR TS="Heptadekafluornonansaeure" OR TS="Perfluornonansaeure" OR TS="Perfluorononanoic acid (PFNA)" OR TS="UNII-5830Z6S63M" OR TS="perfluoro-n-nonanoic acid" OR TS="perfluorononan-1-oic acid" OR TS="perfluorononanoic acid" OR TS="Ammonium Perfluorononanoate" OR TS="Ammonium perfluorononanoate" OR TS="PFNA-H3N") AND PY=2017-2018	2017–2018
Toxline		
Search terms	((pfna OR "c 1800") AND (fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro* OR perfluorinated OR fluorinated OR pfas OR pfos OR pfoa) OR "375-95-1" [rn] OR "2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluorononanoic acid" OR "nonanoic acid 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluoro-" OR "nonanoic acid heptadecafluoro-" OR "perfluoro-n-nonanoic acid" OR "perfluorononan-1-oic acid" OR "perfluorononanoate" OR "perfluorononanoic acid" OR "perfluorononanonic acid" OR "perfluoropelargonic acid" OR "heptadecafluorononanoic acid")) AND (ANEUPL [org] OR BIOSIS [org] OR CIS [org] OR DART [org] OR EMIC [org] OR EPIDEM [org] OR HEEP [org] OR HMTC [org] OR IPA [org] OR RISKLINE [org] OR MTGABS [org] OR NIOSH [org] OR NTIS [org] OR PESTAB [org] OR PPBIB [org]) AND NOT PubMed [org] AND NOT pubdart [org]	No date limit–7/26/2017
Literature update and additional PFNA synonyms search terms	@AND+@OR+(pfna+"c 1800"+fluorocarbon*+"2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluorononanoic+acid"+"nonanoic+acid+2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptadecafluoro-"+"nonanoic+acid+heptadecafluoro-"+"perfluoro-n-nonanoic+acid"+"perfluorononan-1-oic+acid"+perfluorononanoate+"perfluorononanoic+acid"+"perfluoropelargonic+acid"+"heptadecafluorononanoic+acid"+"Methyl-n1-Perfluorononanoic+acid"+"PFNA-n1CH3"+"EINECS 206-801-3"+"Heptadecafluornonansaeure"+"Heptadekafluornonansaeure"+"Perfluornonansaeure"+"Perfluorononanoic+acid (PFNA)"+"UNII-5830Z6S63M"+"perfluoro-n-nonanoic+acid"+"perfluorononan-1-oic+acid"+"perfluorononanoic+acid"+"Ammonium+Perfluorononanoate"+"Ammonium+perfluorononanoate"+"PFNA-H3N"+@TERM+@rn+375-95-1)+@RANGE+yr+2017+2018	2017–2018
TSCATS		
Search terms	"375-95-1" [rn] AND TSCATS[org]	No date limit–7/20/2017

Search	Search strategy	Dates of search
Literature update and additional PFNA synonyms search terms	@TERM+@rn+375-95-1+@RANGE+yr+2017+2018	2017–2018

1

Table B-4. Perfluorohexanoic acid (PFHxA) database search strategy

Search	Search strategy	Dates of search
PubMed		
Search terms	((307-24-4[rn] OR "2,2,3,3,4,4,5,5,6,6,6-undecafluorohexanoic acid"[tw] OR "2,2,3,3,4,4,5,5,6,6,6-undecafluoro-hexanoic acid"[tw] OR "hexanoic acid, 2,2,3,3,4,4,5,5,6,6,6-undecafluoro-"[tw] OR "hexanoic acid, undecafluoro-"[tw] OR "perfluorohexanoic acid"[tw] OR "perfluoro-1-pentanecarboxylic acid"[tw] OR "perfluorocaproic acid"[tw] OR "perfluorohexanoate"[tw] OR "perfluorohexanoic acid"[tw] OR "undecafluoro-1-hexanoic acid"[tw] OR "undecafluorocaproic acid"[tw] OR "undecafluorohexanoic acid"[tw] OR "PFHxA"[tw])) AND ("2017/08/01"[PDAT] : "2018/02/28"[PDAT])	No date limit–7/21/2017
Literature update search terms	((92612-52-7[EC/RN Number]) OR 355-38-4[EC/RN Number]) OR 2062-98-8[EC/RN Number]) OR "PFHxA_ion"[tw]) OR "Perfluorohexanoate"[tw]) OR "Hexanoyl fluoride, 2,2,3,3,4,4,5,5,6,6,6-undecafluoro-"[tw]) OR "Hexanoyl fluoride, undecafluoro-"[tw]) OR "Perfluorohexanoyl fluoride"[tw]) OR "Undecafluorohexanoyl fluoride"[tw]) OR "Perfluoro(2-methyl-3-oxahexanoyl) fluoride"[tw]) OR "Propanoyl fluoride, 2,3,3,3-tetrafluoro-2-(1,1,2,2,3,3,3-heptafluoropropoxy)-" [tw]) OR "Propanoyl fluoride, 2,3,3,3-tetrafluoro-2-(heptafluoropropoxy)-" [tw]) OR "Propionyl fluoride, tetrafluoro-2-(heptafluoropropoxy)-" [tw]) OR "2,2,3,3,4,4,5,5,6,6,6-Undecafluorohexanoic acid"[tw]) OR "EINECS 206-196-6"[tw]) OR "NSC 5213"[tw]) OR "Perfluoro-1-pentanecarboxylic acid"[tw]) OR "Perfluoro-n-hexanoic acid"[tw]) OR "UNII-ZP34Q2220R"[tw]) OR "Undecafluorocaproic acid"[tw]) OR "Ammonium Perfluorohexanoate"[tw]) OR "PFHxA-H3N"[tw]) OR "PFHxA-Na"[tw]) OR "Sodium Perfluorohexanoate"[tw]))	8/1/2017–2/14/2018
Web of Science		
Search terms	((TS="2,2,3,3,4,4,5,5,6,6,6-undecafluorohexanoic acid" OR TS="2,2,3,3,4,4,5,5,6,6,6-undecafluoro-hexanoic acid" OR TS="hexanoic acid, 2,2,3,3,4,4,5,5,6,6,6-undecafluoro-" OR TS="hexanoic acid, undecafluoro-" OR TS="perfluorohexanoic acid" OR TS="perfluoro-1-pentanecarboxylic acid" OR TS="perfluorocaproic acid" OR TS="perfluorohexanoate" OR TS="perfluorohexanoic acid" OR TS="undecafluoro-1-hexanoic acid" OR TS="undecafluorocaproic acid" OR TS="undecafluorohexanoic acid" OR TS="PFHxA")) AND PY=2017-2018	No date limit–7/24/2017

Search	Search strategy	Dates of search
Literature update search terms	TS="PFHxA_ion" OR TS="Perfluorohexanoate" OR TS="Hexanoyl fluoride, 2,2,3,3,4,4,5,5,6,6,6-undecafluoro-" OR TS="Hexanoyl fluoride, undecafluoro-" OR TS="Perfluorohexanoyl fluoride" OR TS="Undecafluorohexanoyl fluoride" OR TS="Perfluoro(2-methyl-3-oxahexanoyl) fluoride" OR TS="Propanoyl fluoride, 2,3,3,3-tetrafluoro-2-(1,1,2,2,3,3,3-heptafluoropropoxy)-" OR TS="Propanoyl fluoride, 2,3,3,3-tetrafluoro-2-(heptafluoropropoxy)-" OR TS="Propionyl fluoride, tetrafluoro-2-(heptafluoropropoxy)-" OR TS="2,2,3,3,4,4,5,5,6,6,6-Undecafluorohexanoic acid" OR TS="EINECS 206-196-6" OR TS="NSC 5213" OR TS="Perfluoro-1-pentanecarboxylic acid" OR TS="Perfluoro-n-hexanoic acid" OR TS="UNII-ZP34Q2220R" OR TS="Undecafluorocaproic acid" OR TS="Undecafluorohexanoic acid" OR TS="Ammonium Perfluorohexanoate" OR TS="PFHxA-H3N" OR TS="PFHxA-Na" OR TS="Sodium Perfluorohexanoate"	2017–2018
Toxline		
Search terms	@AND+@OR+("2,2,3,3,4,4,5,5,6,6,6-undecafluorohexanoic+acid"+"2,2,3,3,4,4,5,5,6,6,6-undecafluoro-hexanoic+acid"+"hexanoic+acid+2,2,3,3,4,4,5,5,6,6,6-undecafluoro-"+"hexanoic+acid+undecafluoro-"+"perfluorohexanoic+acid"+"perfluoro-1-pentanecarboxylic+acid"+"perfluorocaproic+acid"+"perfluorohexanoate"+"perfluorohexanoic acid"+"undecafluoro-1-hexanoic+acid"+"undecafluorocaproic+acid"+"undecafluorohexanoic+acid"+"pfhxa"+@TERM+@rn+(307+24+4)+@RANGE+yr+2017+2018+@NOT+@org+"nih+reporter"	No date limit–7/21/2017
Literature update search terms	@AND+@OR+("PFHxA_ion"+"Perfluorohexanoate"+"Hexanoyl+fluoride,+2,2,3,3,4,4,5,5,6,6,6-undecafluoro-"+"Hexanoyl+fluoride,+undecafluoro-"+"Perfluorohexanoyl+fluoride"+"Undecafluorohexanoyl+fluoride"+"Perfluoro(2-methyl-3-oxahexanoyl)+fluoride"+"Propanoyl+fluoride,+2,3,3,3-tetrafluoro-2-(1,1,2,2,3,3,3-heptafluoropropoxy)-"+"Propanoyl+fluoride,+2,3,3,3-tetrafluoro-2-(heptafluoropropoxy)-"+"Propionyl+fluoride,+tetrafluoro-2-(heptafluoropropoxy)-"+"2,2,3,3,4,4,5,5,6,6,6-Undecafluorohexanoic+acid"+"EINECS+206+196+6"+"NSC+5213"+"Perfluoro-1-pentanecarboxylic+acid"+"Perfluoro-n-hexanoic+acid"+"UNII-ZP34Q2220R"+"Undecafluorocaproic+acid"+"Undecafluorohexanoic+acid"+"Ammonium+Perfluorohexanoate"+"PFHxA-H3N"+"PFHxA-Na"+"Sodium+Perfluorohexanoate")+@NOT+@org+"nih+reporter" @AND+@OR+(@TERM+@rn+92612+52+7+@TERM+@rn+355+38+4+@TERM+@rn+2062+98+8)+@NOT+@org+"nih+reporter"	2017–2018
TSCATS		
Search terms	307-24-4[rn] AND tscats[org]	No date limit–7/20/2017
Literature update search terms	@AND+@OR+(@TERM+@rn+92612+52+7+@TERM+@rn+355+38+4+@TERM+@rn+2062+98+8)+@org+tscat	

1

Table B-5. Perfluorohexanesulfonate (PFHxS) database search strategy

Search	Search strategy	Dates of search
PubMed		
Search terms	108427-53-8[rn] OR 355-46-4[rn] OR "1,1,2,2,3,3,4,4,5,5,6,6,6-Tridecafluorohexane-1-sulfonic acid"[tw] OR "1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-1-Hexanesulfonic acid"[tw] OR "1-Hexanesulfonic acid, 1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-"[tw] OR "1-Hexanesulfonic acid, tridecafluoro-"[tw] OR "1-Perfluorohexanesulfonic acid"[tw] OR "Perfluoro-1-hexanesulfonate"[tw] OR "Perfluorohexane sulfonic acid"[tw] OR "Perfluorohexane-1-sulphonic acid"[tw] OR "Perfluorohexanesulfonate"[tw] OR "Perfluorohexanesulfonic acid"[tw] OR "Perfluorohexylsulfonate"[tw] OR "Tridecafluorohexanesulfonic acid"[tw] OR "tridecafluoro-1-Hexanesulfonic acid"[tw] OR "PFHxS"[tw]	No date limit–7/21/2017
Literature update search terms	((108427-53-8[EC/RN Number]) OR 423-50-7[EC/RN Number]) OR "1-Hexanesulfonic acid, 1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-, ion(1-)"[tw]) OR "PFHxS ion(1-)"[tw]) OR "PFHxS_ion"[tw]) OR "Perfluorohexanesulfonate"[tw]) OR "Tridecafluorohexane-1-sulfonate"[tw]) OR "perfluorohexyl sulfonate"[tw]) OR "1,1,2,2,3,3,4,4,5,5,6,6,6-Tridecafluoro-1-hexanesulfonyl fluoride"[tw]) OR "1-Hexanesulfonyl fluoride, 1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-"[tw]) OR "1,1,2,2,3,3,4,4,5,5,6,6,6-Tridecafluoro-1-hexanesulfonic acid"[tw]) OR "EC 206-587-1"[tw]) OR "EINECS 206-587-1"[tw]) OR "PFHS"[tw]) OR "Perfluorhexan-1-sulfonsaure"[tw]) OR "Perfluorohexane sulfonic acid (PFHxS)"[tw]) OR "Perfluorohexane-1-sulphonic acid"[tw]) OR "acide perfluorohexane-1-sulfonique"[tw]) OR "acido perfluorohexano-1-sulfonico"[tw]) OR "perfluorohexane-1-sulphonic acid"[tw]) OR "perfluorohexanesulfonic acid"[tw]) OR "Ammonium Perfluorohexanesulfonate"[tw]) OR "Ammonium perfluorohexanesulfonate"[tw]) OR "PFHxS-H3N"[tw]) OR "PFHxS-K"[tw]) OR "Potassium Perfluorohexanesulfonate"[tw]) OR "Potassium perfluorohexanesulfonate"[tw]) OR "Lithium Perfluorohexanesulfonate"[tw]) OR "Lithium perfluorohexanesulfonate"[tw]) OR "PFHxS-Li"[tw]))	8/1/2017–2/14/2018
Web of Science		
Search terms	TS="1,1,2,2,3,3,4,4,5,5,6,6,6-Tridecafluorohexane-1-sulfonic acid" OR TS="1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-1-Hexanesulfonic acid" OR TS="1-Hexanesulfonic acid, 1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-" OR TS="1-Hexanesulfonic acid, tridecafluoro-" OR TS="1-Perfluorohexanesulfonic acid" OR TS="Perfluoro-1-hexanesulfonate" OR TS="Perfluorohexane sulfonic acid" OR TS="Perfluorohexane-1-sulphonic acid" OR TS="Perfluorohexanesulfonate" OR TS="Perfluorohexanesulfonic acid" OR TS="Perfluorohexylsulfonate" OR TS="Tridecafluorohexanesulfonic acid" OR TS="tridecafluoro-1-Hexanesulfonic acid" OR TS="PFHxS"	No date limit–7/24/2017

Search	Search strategy	Dates of search
Literature update search terms	TS="1-Hexanesulfonic acid, 1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-, ion(1-)" OR TS="PFHxS ion(1-)" OR TS="PFHxS_ion" OR TS="Perfluorohexanesulfonate" OR TS="Tridecafluorohexane-1-sulfonate" OR TS="perfluorohexyl sulfonate" OR TS="1,1,2,2,3,3,4,4,5,5,6,6,6-Tridecafluoro-1-hexanesulfonyl fluoride" OR TS="1-Hexanesulfonyl fluoride, 1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-" OR TS="1,1,2,2,3,3,4,4,5,5,6,6,6-Tridecafluoro-1-hexanesulfonic acid" OR TS="EC 206-587-1" OR TS="EINECS 206-587-1" OR TS="PFHS" OR TS="Perfluorhexan-1-sulfonsaure" OR TS="Perfluorohexane sulfonic acid (PFHxS)" OR TS="Perfluorohexane-1-sulphonic acid" OR TS="acide perfluorohexane-1-sulfonique" OR TS="acido perfluorohexano-1-sulfonico" OR TS="perfluorohexane-1-sulphonic acid" OR TS="perfluorohexanesulfonic acid" OR TS="Ammonium Perfluorohexanesulfonate" OR TS="Ammonium perfluorohexanesulfonate" OR TS="PFHxS-H3N" OR TS="PFHxS-K" OR TS="Potassium Perfluorohexanesulfonate" OR TS="Potassium perfluorohexanesulfonate" OR TS="Lithium Perfluorohexanesulfonate" OR TS="Lithium perfluorohexanesulfonate" OR TS="PFHxS-Li"	2017–2018
Toxline		
Search terms	(108427-53-8[rn] OR 355-46-4[rn] OR "1,1,2,2,3,3,4,4,5,5,6,6,6-Tridecafluorohexane-1-sulfonic acid" OR "1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-1-Hexanesulfonic acid" OR "1-Hexanesulfonic acid, 1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-" OR "1-Hexanesulfonic acid, tridecafluoro-" OR "1-Perfluorohexanesulfonic acid" OR "Perfluoro-1-hexanesulfonate" OR "Perfluorohexane sulfonic acid" OR "Perfluorohexane-1-sulphonic acid" OR "Perfluorohexanesulfonate" OR "Perfluorohexanesulfonic acid" OR "Perfluorohexylsulfonate" OR "Tridecafluorohexanesulfonic acid" OR "tridecafluoro-1-Hexanesulfonic acid" OR "PFHxS") AND (ANEUPL [org] OR BIOSIS [org] OR CIS [org] OR DART [org] OR EMIC [org] OR EPIDEM [org] OR HEEP [org] OR HMTC [org] OR IPA [org] OR RISKLINE [org] OR MTGABS [org] OR NIOSH [org] OR NTIS [org] OR PESTAB [org] OR PPBIB [org]) [not] PubMed [org] [not] pubdart [org]	No date limit–7/21/2017
Literature update search terms	@AND+@OR+("1-Hexanesulfonic+acid,+1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-,+ion(1-)"+"PFHxS+ion(1-)"+"PFHxS_ion"+"Perfluorohexanesulfonate"+"Tridecafluorohexane-1-sulfonate"+"perfluorohexyl+sulfonate"+"1,1,2,2,3,3,4,4,5,5,6,6,6-Tridecafluoro-1-hexanesulfonyl+fluoride"+"1-Hexanesulfonyl+fluoride,+1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-"+"1,1,2,2,3,3,4,4,5,5,6,6,6-Tridecafluoro-1-hexanesulfonic+acid"+"EC+206-587-1"+"EINECS+206-587-1"+"PFHS"+"Perfluorhexan-1-sulfonsaure"+"Perfluorohexane+sulfonic+acid+(PFHxS)"+"Perfluorohexane-1-sulphonic+acid"+"acide+perfluorohexane-1-sulfonique"+"acido+perfluorohexano-1-sulfonico"+"perfluorohexane-1-sulphonic+acid"+"perfluorohexanesulfonic+acid"+"Ammonium+Perfluorohexanesulfonate"+"Ammonium+perfluorohexanesulfonate"+"PFHxS-H3N"+"PFHxS-K"+"Potassium+Perfluorohexanesulfonate"+"Potassium+perfluorohexanesulfonate"+"Lithium+Perfluorohexanesulfonate"+"Lithium+perfluorohexanesulfonate"+"PFHxS-Li")+@NOT+@org+"nih+reporter" @OR+(@TERM+@rn+108427+53+8+@TERM+@rn+423+50+7)+@NOT+@org+"nih+reporter"	2017–2018

Search	Search strategy	Dates of search
TSCATS		
Search terms	@OR+(@term+@rn+355-46-4+@term+@rn+108427-53-8) +@AND+@org+tscats	No date limit−7/21/2017
Literature update search terms	@OR+(@TERM+@rn+"108427+53+8"+@TERM+@rn+"423+50+7")+@org+tscats	2017−2018

1

Table B-6. Title/abstract-level screening criteria for the initial literature searches

	Inclusion criteria	Exclusion criteria
Populations	• Humans • Standard mammalian animal models, including rat, mouse, rabbit, guinea pig, hamster, monkey, dog • Alternative animal models in standard laboratory conditions (e.g., Xenopus, zebrafish, minipig) • Human or animal cells, tissues, or organs (not whole animals); bacteria, nonmammalian eukaryotes; other nonmammalian laboratory species	• Ecological species
Exposures	• Exposure is to a PFAS compound • Exposure via oral, inhalation, dermal, intraperitoneal, or intravenous injection routes • Exposure is measured in air, dust, drinking water, diet, gavage, injection or via a biomarker of exposure (PFAS levels in whole blood, serum, plasma, or breastmilk)	• Study population is not exposed to a PFAS compound • Exposure is to a mixture only
Outcomes	• Studies that include a measure of one or more health effect endpoints, including but not limited to, effects on reproduction, development, developmental neurotoxicity, liver, thyroid, immune system, nervous system, genotoxicity, and cancer • *In vivo* and/or *in vitro* studies related to toxicity mechanisms, physiological effects/adverse outcomes, and studies useful for elucidating toxic modes of action (MOAs) • Qualitative or quantitative description of absorption, distribution, metabolism, excretion, toxicokinetic and/or toxicodynamic models (e.g., PBPK, PBTK, PBTK/TD) • Studies addressing risks to infants, children, pregnant women, occupational workers, the elderly, and any other susceptible or differentially exposed populations	

	Inclusion criteria	Exclusion criteria
Other	• Structure and physiochemical properties • Reviews and regulatory documents	• Not on topic, including: • Abstract only, inadequately reported abstract, or no abstract and not considered further because study was not potentially relevant • Bioremediation, biodegradation, or chemical or physical treatment of PFAS compounds, including evaluation of wastewater treatment technologies and methods for remediation or contaminated water and soil • Ecosystem effects • Studies of environmental fate and transport of PFAS compounds in environmental media • Analytical methods for detecting/measuring PFAS compounds in environmental media and use in sample preparations and assays • Studies describing the manufacture and use of PFAS compounds • Not chemical specific (studies that do not involve testing of PFAS compounds) • Studies that describe measures of exposure to PFAS compounds without data on associated health effects

MOA = mode of action; PBPK = physiologically based pharmacokinetic; PBTK = physiologically based toxicokinetic; TD = toxicodynamic; PFAS = per- and polyfluoroalkyl substances.

1

Table B-7. Example DistillerSR form questions to be used for title/abstract- and full text-level screening for literature search updates from 2019

	Used in title/abstract and full-text screening					Used in full text only	
Question	**Source of study if not identified from database search?**	**Does the article meet PECO criteria?**	**If meets PECO, what type of evidence?**	**If supplemental, what type of information?**	**Which PFAS did the study report?**	**If meets PECO, which health outcome(s) apply?**	**If meets PECO and endocrine outcome, which endocrine tags apply?**
Answer options (can select multiple options)	• Source other than HERO database search	• Yes • No • Unclear • Tag as potentially relevant supplemental information	• Human • Animal (mammalian models) • *In vitro* or *in silico* genotoxicity • PBPK or PK model	• *In vivo* mechanistic or MOA studies, including non-PECO routes of exposure (e.g., injection) and populations (e.g., nonmammalian) • *In vitro* or *in silico* studies (non-genotoxicity) • ADME/ toxicokinetic (excluding models) • Exposure assessment or characterization (no health outcome) • PFAS Mixture Study (no individual PFAS comparisons) • Human case reports or case series • Ecotoxicity studies	• PFBA • PFHxA • PFHxS • PFNA • PFDA	• General toxicity, including body weight, mortality, and survival • Cancer • Cardiovascular, including serum lipids • Endocrine (hormone) • Gastrointestinal • Genotoxicity • Growth (early life) and development • Hematological, including non-immune/hepatic/ renal clinical chemistry measures • Hepatic, including liver measures and serum markers (e.g., ALT; AST) • Immune/Inflammation • Musculoskeletal	• Adrenal • Sex hormones (e.g., androgen; estrogen; progesterone) • Neuroendocrine • Pituitary • Steroidogenesis • Thyroid

				• Environmental fate or occurrence (including food) • Manufacture, engineering, use, treatment, remediation, or laboratory methods • Other assessments or records with no original data (e.g., reviews, editorials, commentaries)		• Nervous system, including behavior and sensory function • Nutrition and metabolic • Ocular • PBPK or PK model • Renal, including urinary measures (e.g., protein) • Reproductive • Respiratory • Skin and connective tissue effects	

ADME = absorption, distribution, metabolism, and excretion; ALT = alanine aminotransferase; AST = aspartate transaminase; HERO = Health and Environmental Research Online; MOA = mode of action; PBPK = physiologically based pharmacokinetic; PECO = populations, exposures, comparators, and outcomes; PFAS = per- and polyfluoroalkyl substances; PFBA = perfluorobutanoic acid; PFDA = perfluorodecanoic acid; PFHxA = perfluorohexanoic acid; PFHxS = perfluorohexanesulfonate; PFNA = Perfluorononanoic acid; PK = pharmacokinetic.

1