

Methods for range finding experiments and reproductive and developmental toxicity tests with *Peromyscus* exposed to per- and polyfluoroalkyl substances (PFAS)

Prepared by: Allison M. Narizzano

U.S. Army Public Health Center, United States Army, Aberdeen Proving Ground, Maryland, United States of America

For more information:

usarmy.apg.medcom-aphc.mbx.tox-info@mail.mil

Approved for public release; distribution unlimited.

The views expressed in this paper are those of the author(s) and do not necessarily reflect the official policy of the Department of Defense, Department of the Army, U.S. Army Medical Department or the U.S. Government.

October 2019



TOXICOLOGICAL STUDY NO. S.0043781-18 AND S.0059303-19
PROTOCOL NO. 67-17-04-01 AND 03-18-08-01
DATA TO SUPPORT THE DEVELOPMENT OF WILDLIFE TOXICITY REFERENCE
VALUES FOR PER- AND POLYFLUOROALKYL SUBSTANCES
OCTOBER 2017–AUGUST 2019

1. SUMMARY

Per- and polyfluoroalkyl substances (PFAS) are compounds manufactured for use in multiple products to include paints, cleaning agents, fire suppressants, nonstick cookware, food containers, and water-impermeable products in general. PFAS are also key ingredients in highly efficient fire suppressant agents known as Aqueous Film Forming Foams (AFFFs), which have been used by the Department of Defense (DOD) since approximately 1970 for fire training and emergency response activities.

Concerns about PFAS stem from their ubiquitous presence in the environment, widespread reports of toxicity, and the resistance of these compounds to degradation. The primary goal of the range finding experiments was to ensure bioaccumulation and sufficient body burden of the chemicals for reproductive and developmental toxicity tests. Data from the reproductive and developmental toxicity tests will be used to derive Toxicity Reference Values (TRVs), which are critical components of environmental risk assessments that help determine if the risk of environmental exposure is acceptable. This study describes the methods used to develop data to be used to develop TRVs; data collection and evaluation are underway.

In the range finding experiments, PFAS were administered orally to white-footed mice (*Peromyscus leucopus*) for 28 consecutive days. Blood samples were collected every 7 days via facial/submandibular venipuncture and analyzed for the concentration of PFAS. At the end of the study, selected tissues were weighed and processed for histopathology. This study provided information on the target organs of individual PFAS and the internal dosimetry of each compound. These data were used to determine dose levels and refine the design for reproductive and developmental toxicity studies with individual PFAS.

In the reproductive and developmental toxicity tests, PFAS were administered orally to white-footed mice for 28 consecutive days, after which mating pairs were established. Daily oral exposure to PFAS continued until a litter was generated, until 12 weeks of co-housing had elapsed, or 16 weeks of co-housing elapsed without the generation of a litter, depending on the experiment. Fetuses and offspring were exposed to the test compounds through pregnancy and lactation. All animals were monitored for body weight changes and signs of toxicity. Animals were sensitized with sheep red blood cells (SRBCs) toward the end of the dosing, and immune function was evaluated via the Jerne plaque forming assay immediately following euthanasia. Other samples collected at termination were blood for analytical chemistry and organs for histological analysis. Appropriate data from this study will be used to derive TRVs. The most important criterion for inclusion in TRV derivation is that the toxic effects are most likely to influence population stability (e.g., mortality, reproduction, development, growth, immunotoxicity, or behavior relevant to reproduction, feeding, and predator avoidance).

2. RANGE FINDING EXPERIMENTS

2.1 Animals and Housing Conditions

Adult male (n=196; 22.58 ± 0.27 g) and female (n=198; 20.49 ± 0.21 g) *Peromyscus leucopus* were obtained from the *Peromyscus* Genetic Stock Center at the University of South Carolina. Animals were obtained for Experiments 1 and 2 in one shipment and they were obtained for Experiment 3 in a separate shipment. Animals were acclimated to the facility for at least 5 days before initiation of dosing in Experiments 1 and 3; they were acclimated to the facility for 5 months before initiating dosing in Experiment 2. Experiments 1 and 2 were supposed to be investigated concurrently; however, unexpected solubility issues necessitated delaying the investigation of one chemical and proceeding with investigation of three chemicals.

The following were the age ranges for male and female mice at the initiation of each experiment:

- Experiment 1: males were between 133 and 222 days; females were between 136 and 238 days.
- Experiment 2: males were between 301 and 390 days; females were between 304 and 406 days.
- Experiment 3: males were between 343 and 392 days; females were between 301 and 392 days.

The age range of the animals is partially due to the delay of Experiment 2. However this is also because there is one commercial vendor for the *P. leucopus*, breeding is seasonal and sporadic, and litter sizes are smaller (roughly 2.7 pups/litter) than those of traditional laboratory mice (roughly 8 pups/litter) (Schmidt 2012; Plumel 2014).

Assignment to a dose group and chemical was accomplished using a stratified random procedure, with animals stratified according to pre-study body weight and groups assigned by random draw. Within each experiment, body weight did not differ among groups (i.e., chemical or dose level) prior to initiation of dosing. A cage card uniquely identified each animal. Within each cage, animals were identified by a tail color.

All animals were housed in temperature-, relative humidity-, and light-controlled rooms. The target conditions of the rooms were 68–72 degrees Fahrenheit (°F) and 30–70% humidity. An automatically controlled 12/12-hour light/dark cycle was maintained, with the dark period beginning at 1800 hours. A certified pesticide-free rodent chow (Harlan Teklad® 2016C Certified Rodent Diet) was available *ad libitum*. Depending on the experiment, animals were provided with filtered tap water or treated filtered tap water *ad libitum*. Animals were same sex group housed in suspended 7" wide x 10.5" tall x 5" deep plastic cages on a static, stainless steel rack. The U.S. Army Public Health Center (APHC) Institutional Animal Care and Use Committee approved the animal care and use procedures. Animal care and use was conducted according to "The Guide for the Care and Use of Laboratory Animals" and all applicable Federal and DOD regulations (NRC 2011). The APHC Animal Care and Use Program is fully accredited by the AAALAC International. This study was conducted according to Good Laboratory Practices (CFR 2011).

2.2 Test Substances

The six test substances (see Table 1) were obtained from Sigma Aldrich, St. Louis, Missouri. The manufacturer conducted purity analyses for these compounds. The six test substances are—

- Heptadecafluorooctanesulfonic acid potassium salt (PFOS),
- Perfluorooctanoic acid (PFOA),
- Tridecafluorohexane-1-sulfonic acid potassium salt (PFHxS),
- Potassium nonafluoro-1-butanefulfonate (PFBS),
- Tridecafluorooctane-1-sulphonic acid (6:2 FTS), and
- Perfluorononanoic acid (PFNA).

Table 1. Test Substances

Chemical	Abbreviation	Molecular weight	CAS-No.	Purity
Heptadecafluorooctanesulfonic acid potassium salt	PFOS	538.22 g/mol	2795-39-3	≥98%
Perfluorooctanoic acid	PFOA	414.07 g/mol	335-67-1	>95%
Tridecafluorohexane-1-sulfonic acid potassium salt	PFHxS	438.20 g/mol	3871-99-6	≥98%
Potassium nonafluoro-1-butanefulfonate	PFBS	338.19 g/mol	29420-49-3	98%
Tridecafluorooctane-1-sulphonic acid	6:2 FTS	428.17 g/mol	27619-97-2	98%
Perfluorononanoic acid	PFNA	464.08 g/mol	375-95-1	97%

2.3 Dose Selection and Test Substance Preparation and Administration

Dose selection was based on the ultimate objective of the study, which was to identify possible target organs of individual PFAS and the internal dosimetry of each compound. The intent was not to cause or detect toxicity, if possible.

2.3.1 Experiment 1

Dosing solutions were prepared by—

- Weighing the required amount of neat PFOA, PFHxS, and PFBS,
- Transferring to a 1,000 milliliters (mL) volumetric flask,
- Adding approximately 900 mL of animal drinking water,
- Stirring using a magnetic stir bar and stir plate until dissolved, and
- Adding water to the 1,000 mL mark.

Three drinking water dosing solutions (37.5, 75, and 150 milligrams per milliliter (mg/L)) were used through the experiment. Using an assumed default water consumption rate of 0.003 liters per kilogram per day (L/kg-day) and a default body weight of 0.0225 kg, dosing solutions

corresponded to approximately 5, 10, and 20 milligrams per kilogram per day (mg/kg-day). Solutions were prepared weekly. Drinking water/dosing reservoirs were replaced completely every 4 days and reservoirs were changed every 2 weeks.

2.3.2 Experiment 2.1

Dosed treats were prepared by—

- Weighing the required volume of neat PFOS, and
- Adding the required amount of methanol to make a stock solution (18.5 mg/mL).

Stock solution was applied to Fruit Loops™ with a pipette (at variable volumes) to create dosed treats. Treats were dried for 24 hours to ensure methanol evaporated off prior to being administered to animals. Three doses, 2.5, 25, and 250 mg/kg-day, were used through the experiment. Stock solution was prepared daily. Treats were prepared daily, 1 day prior to administration.

2.3.3 Experiment 2.2

Dosing suspensions were prepared by—

- Weighing the required amount of neat PFOS,
- Transferring to a polypropylene container,
- Measuring the appropriate volume of filtered tap water or corn oil using a graduated cylinder, and
- Adding the tap water or corn oil to the polypropylene container.

Three dosing solutions or suspensions, 0.25, 4.0, and 10.0 mg/mL, were used throughout the experiment, which corresponded to 2.5, 40, and 100 mg/kg-day. Solutions were prepared once and suspensions were prepared bi-weekly.

2.3.4 Experiment 3

Dosing solutions and suspensions were prepared by—

- Weighing the required volume of neat 6:2 FTS and PFNA,
- Adding the required amount of methanol to make a stock solution (500 mg/mL),
- Adding the appropriate amount of stock solution to a 50 mL volumetric flask,
- Adding approximately 40 mL of water from the animal room,
- Stirring using a magnetic stir bar and stir plate until dissolved, and
- Adding water to the 50 mL mark.

Solutions were prepared bi-weekly. Three dosing solutions, 0.25, 1.0, and 2.0 mg/mL, were prepared for 6:2 FTS, which corresponded to 2.5, 10, and 20 mg/kg-day. The 0.25 and 1.0 mg/mL 6:2 FTS dosing solutions were volumetrically adjusted such that they had the same volume of methanol as the 2.0 mg/mL 6:2 FTS dosing solution.

Toxicological Study No. S.0043781-18 and S.0059303-19, October 2017–August 2019

Two dosing solutions, 0.25 and 1.0 mg/mL, were prepared for PFNA, which corresponded to 2.5 and 10 mg/kg-day. One dosing suspension, 2.0 mg/mL, was prepared for PFNA, which corresponded to 20 mg/kg-day. The 0.25 and 1.0 mg/mL PFNA dosing solutions were volumetrically adjusted such that they had the same volume of methanol as the 2.0 mg/mL PFNA dosing suspension.

Dosing solutions were prepared by—

- Weighing the required volume of neat PFHxS,
- Transferring to a 50 mL volumetric flask,
- Adding approximately 40 mL of water from the animal room,
- Stirring using a magnetic stir bar and stir plate until dissolved, and
- Adding water to the 50 mL mark.

Two dosing solutions, 1.0 and 2.0 mg/mL, were used through the experiment, which corresponded to 10 and 20 mg/kg-day.

Because of the shared control group with 6:2 FTS and PFNA, the 1.0 and 2.0 mg/mL dosing solutions were volumetrically adjusted such that they had the same volume of methanol as the 2.0 mg/mL (6:2 FTS and PFNA) dosing solution and suspension.

2.4 Study Design

2.4.1 Experiment 1

Male and female white-footed mice were exposed to PFOA (9/sex/dose), PFHxS (9/sex/dose), and PFBS (8–9/sex/dose) via drinking water for 28 consecutive days. This experiment utilized a shared control group.

2.4.2 Experiment 2.1

Male and female white-footed mice were exposed to PFOS (8/sex/dose) via food treats 2 and 3 days, respectively. Food treats are a well-documented exposure route for toxicity testing with rodents, including flame retardant materials (Patisaul et al. 2009; McCaffrey et al. 2013; Hays et al. 2002)

2.4.3 Experiment 2.2

Male and female white-footed mice were exposed to PFOS (8–9/sex/dose) via oral gavage for 28 consecutive days. Male and female mice received their first four and three doses, respectively, via filtered tap water. Animals received their remaining doses via corn oil (i.e., days 5 and 4 through 28 for males and females, respectively).

2.4.4 Experiment 3

Male and female white-footed mice were exposed to 6:2 FTS (9/sex/dose) and PFNA (9/sex/dose) via oral gavage for 28 consecutive days. Male and female white-footed mice were also exposed to PFHxS (6-9/sex/dose) via oral gavage for 28 consecutive days. Because PFHxS was administered to white-footed mice via drinking water in Experiment 1, fewer animals were used in Experiment 3. This experiment utilized a shared control group.

2.5 Clinical Observations and Body Weight

Animals were removed from their home cages and observed daily by study personnel in conjunction with dosing for signs of toxicity, morbidity, and mortality. Body weights were taken at the start of test substance administration, at least weekly thereafter, and at termination.

2.6 Blood Collection

Blood was collected via submandibular venipuncture on days 0, 7, 14, and 21 of test substance administration. Blood was collected in weigh boats free of additives, transferred to microcentrifuge tubes, allowed to clot at room temperature for 60 to 120 minutes, and centrifuged twice for approximately 10 minutes at 600 x g. Serum was removed and frozen at -80 degrees Celsius (°C).

2.7 Necropsy, Trunk Blood Collection, and Organ Mass

After 28 days of dosing, surviving animals were rendered unconscious with carbon dioxide (CO₂), euthanized via decapitation, and trunk blood was collected for serum analysis.

A full, detailed gross necropsy, including a careful examination of the external surface of the body, all orifices, and the cranial, thoracic, and abdominal cavities and their contents, was performed on all experimental animals following euthanasia. At necropsy, the brain, heart, kidneys, liver, ovaries, uterus, spleen, thymus, testes, and epididymides were removed, trimmed, and weighed. Testes were weighed as pairs. Any observed lesions were retained for processing.

The brain, heart, kidneys, liver, ovaries, uterus, spleen, and thymus were stored in 10% buffered formalin for fixation. The testes and right epididymis from each animal were placed in Davidson's fixative overnight (no longer than 24 hours). After fixation, the tissues were rinsed with deionized water and stored in 70% ethanol.

2.8 Sperm Analysis

Sperm analysis was conducted for Experiment 3. The left epididymis from each male was removed during necropsy and submitted for sperm analysis. In preparation, 10 mL of Gibco® Medium 199 (M199) was pipetted into each well of Corning® Costar® 6-well cell culture plates and warmed on a slide warmer to approximately 37°C. Each epididymis was weighed, placed in a well containing M199, minced using small scissors, and incubated for 5 minutes at

approximately 37°C. Then, samples were loaded on a standard count chamber slide (Leja®). Slides were immediately loaded into a Hamilton-Thorne IVOS Sperm Analysis System® and read using the IDENT® program.

2.9 Histopathology

Preserved tissues were processed using standard techniques, sectioned 4–5 micrometers (µm) thick, and stained with hematoxylin and eosin using a routine automatic stainer. Testis and epididymis were additionally stained with periodic acid-Schiff (PAS) stain. Histopathology was performed for all high-dose and control animals at the discretion of the pathologist and based on observed toxicity and gross pathology findings. Lower dose groups were examined if exposure-related effects were seen in the high-dose group, if gross lesions were present, or if other signs of organ toxicity were noted (e.g., changes in organ mass). Histopathologic findings were subjectively graded across the following 6-point scale:

- Grade 0 (essentially normal) referred to tissues with changes observed in <1% of the sampled tissue.
- Grade 1 (minimum) referred to a change which affected <5% of the presented tissue area.
- Grade 2 (mild) referred to a change, which affected 6 to 20% of the tissue area.
- Grade 3 (moderate) was scaled to refer to a change, which affected 21 to 40% of the tissue area.
- Grade 4 (marked) was scaled for lesions affecting 41 to 80% of the tissue area.
- Grade 5 (severe) indicated that >80% of the tissue was affected (Schafer et al. 2018).

2.10 Serum Analysis

Serum analysis (Day 0, 21, and 28) was conducted at the Colorado School of Mines based on published methods (Reiner et al. 2009). Serum analysis (Day 28) was conducted at a second lab (3M) for inter-lab verification purposes based on published methods (Sundstrom et al. 2012).

2.11 Data Collection and Statistical Analyses

Experimental data generated during the course of this study were recorded by hand and tabulated, summarized, and/or statistically analyzed using Microsoft® Excel, Minitab®, or SPSS®. Environmental data were automatically recorded using MetaSys® Building Management System.

Data not normally distributed were log transformed and retested for normality. Variance equality was determined by Levene's test. Analyses were conducted for male and female mice separately. Parameters measured multiple times (i.e., body weight) were analyzed using repeated one-way analysis of variance (ANOVA), and those measured at the end of the study (i.e., organ weights) were analyzed using ANOVA with dose group as the main effect. Absolute organ mass was analyzed by analysis of covariance (ANCOVA), using dose group as the main effect and body weight at necropsy as the covariate (Bailey et al. 2004). Fisher's exact test was used to determine significant differences between treated and control groups for nominal or count data (e.g., histology, litter incidence, etc.). Histopathology severity categories were

consolidated to present/absent for statistical analyses. Significant overall effects were followed with comparisons between the control group and each treatment group. Statistical significance are defined at the $p < 0.05$ level. If dose effect was significant ($p < 0.05$), appropriate post hoc analyses were performed (Tukey's multiple comparison (ANOVA) or Sidak (for ANCOVA). Interpretation of changes in absolute organ mass, organ-to-body mass ratio, and organ-to-brain mass ratio in the evaluation of compound-related effects was based on published analysis of control animal data (Bailey et al. 2004).

3. REPRODUCTIVE AND DEVELOPMENTAL TOXICITY TESTS

3.1 Animals and Housing Conditions

For Experiment 1, adult male ($n=124$; 22.45 ± 0.31 g) and female ($n=124$; 19.80 ± 0.26 g) *Peromyscus leucopus* were obtained from the Peromyscus Genetic Stock Center at the University of South Carolina. For Experiment 2, adult male ($n=125$; 21.23 ± 0.25 g) and female ($n=125$; 18.90 ± 0.23 g) *Peromyscus leucopus* were obtained from the Peromyscus Genetic Stock Center at the University of South Carolina. Animals were acclimated to the facility for at least 5 weeks before initiation of dosing in Experiments 1 and 2.

The following were the age ranges for male and female mice at the initiation of each experiment:

- Experiment 1: males were between 150 and 412 days; females were between 203 and 406 days.
- Experiment 2: males were between 118 and 349 days; females were between 117 and 357 days.

The age range of the animals varies because there is one commercial vendor for the *P. leucopus*, breeding is seasonal and sporadic, and litter sizes are smaller (roughly 2.7 pups/litter) than those of traditional laboratory mice (roughly 8 pups/litter) (Schmidt 2012; Plumel 2014).

In Experiment 1, assignment to a dose group was accomplished using a stratified random procedure, with animals stratified according to pre-study body weight and evenly spread among the dose groups, as well as by taking age and known genetic background into consideration. In Experiment 1, mating pairs were randomly assigned within each dose group. In Experiment 2, assignment to a dose group was accomplished using a stratified random procedure, with animals stratified according to age and evenly spread among the dose groups, as well as by taking pre-study body weight and known genetic background into consideration. In Experiment 2, mating pairs were assigned according to age. Within each experiment, body weight did not differ among dose groups prior to initiation of dosing. A cage card uniquely identified each animal. Within each cage, animals were identified by a tail color.

All animals were housed in temperature-, relative humidity-, and light-controlled rooms. The target conditions of the rooms were 68–72°F and 30–70% humidity. An automatically controlled 12/12-hour light/dark cycle was maintained for the first half of Experiment 1, with the dark period beginning at 1800 hours. For the second half of Experiment 1 and for all of Experiment 2, an

automatically controlled 16/8-hour light/dark cycle was maintained, with the dark period beginning at 2000 hours. The appropriate light cycle for *Peromyscus* breeding is a 16/8-hour light/dark cycle (Crossland 2006). A certified pesticide-free rodent chow (Harlan Teklad 2016C Certified Rodent Diet) was available *ad libitum*. Animals were same sex group housed in suspended 7" wide x 10.5" tall x 5" deep plastic cages on a static, stainless steel rack. The APHC Institutional Animal Care and Use Committee approved animal care and use procedures. Animal care and use was conducted according to "The Guide for the Care and Use of Laboratory Animals" and all applicable Federal and DOD regulations (NRC 2011). The APHC Animal Care and Use Program is fully accredited by the AAALAC International. This study was conducted consistent with Good Laboratory Practices (CFR 2011).

3.2 Test Substances

Test substances were obtained from Sigma Aldrich, St. Louis, Missouri. Purity analyses for these compounds were conducted by the manufacturer (see Table 1).

3.3 Dose Selection and Test Substance Preparation and Administration

3.3.1 Experiment 1

Dosing suspensions were prepared by—

- Weighing the required amount of neat PFOS,
- Transferring to a polypropylene container,
- Measuring the appropriate volume of filtered tap water using a graduated cylinder, and
- Adding the tap water to the polypropylene container.

Three dosing solutions 0.02, 0.1, and 0.5 mg/mL, were used throughout the experiment, which corresponded to 0.2, 1.0, and 5.0 mg/kg-day. Solutions were prepared bi-weekly.

3.3.2 Experiment 2

Dosing solutions were prepared by—

- Weighing the required volume of neat PFNA,
- Adding the required amount of methanol to make a stock solution (500 mg/mL),
- Adding the appropriate amount of stock solution to a 1,000 mL volumetric flask,
- Adding approximately 600 mL of water from the animal room,
- Stirring using a magnetic stir bar and stir plate until dissolved, and
- Adding water to 400 mL.

Four dosing solutions, 0.004, 0.02, and 0.1, and 0.5 mg/mL, were used throughout the experiment, which corresponded to 0.04, 0.2, 1.0, and 5.0 mg/kg-day. Solutions were prepared bi-weekly.

3.4 Study Design

3.4.1 Experiment 1

PFOS was administered via oral gavage to male and female Parental (P) generation animals for 28 consecutive days prior to establishing mating pairs. Pairs were orally exposed until a litter was generated or until 16 weeks of co-housing had elapsed without the generation of a litter. Fetuses and offspring (First Filial (F1) generation animals) were exposed to PFOS through pregnancy and lactation. F1 generation animals were maintained until postnatal day (PND) 57–63 to batch the animals in manageable groups.

3.4.2 Experiment 2

PFNA was administered via oral gavage to male and female P generation animals for 28 consecutive days prior to establishing mating pairs. Pairs were orally exposed for a total of 16 weeks, regardless of generation of a litter. Five days prior to euthanasia, P generation animals were sensitized by intraperitoneal injection with 25% SRBC (Lampire Biological Labs, Pipersville, Pennsylvania) in a phosphate buffered solution (PBS). Fetuses and offspring (F1 generation animals) were exposed to PFNA through pregnancy and lactation. F1 generation animals were culled at PND 10.

3.5 Clinical Observations and Body Weight

Animals were removed from their home cages and observed daily by study personnel in conjunction with dosing for signs of toxicity, morbidity, and mortality. Body weights were taken at the start of test substance administration, at least weekly thereafter, and at termination.

3.6 Necropsy, Trunk Blood Collection, and Organ Mass

In Experiment 1, after weaning a litter at PND 25 (or after 16 weeks of co-housing without the generation of a litter), surviving P generation animals were rendered unconscious with CO₂, euthanized via decapitation, and trunk blood was collected for hormone analysis (see paragraph 3.10). In Experiment 1, at PND 57–63, surviving F1 generation animals were rendered unconscious with CO₂, euthanized via decapitation, and trunk blood was collected for hormone analysis (if possible).

In Experiment 2, surviving P generation animals were rendered unconscious with CO₂ after 16 total weeks of exposure to PFNA, euthanized via decapitation, and trunk blood was collected for hormone analysis. In Experiment 2, at PND 10, surviving F1 generation animals were rendered unconscious with CO₂ and euthanized via decapitation.

A full, detailed gross necropsy, including a careful examination of the external surface of the body, all orifices, and the cranial, thoracic, and abdominal cavities and their contents, was performed on all animals in Experiment 1 and all P generation animals in Experiment 2 following euthanasia. At necropsy, the brain, heart, kidneys, liver, ovaries, uterus, spleen, thymus, testes, and epididymides were removed, trimmed, and weighed. Testes were weighed as pairs. Any observed lesions were retained for processing. In Experiments 1 and 2, one epididymis was

retained for histopathology and one was collected for sperm analysis (see paragraph 3.7). In Experiment 2, spleens were aseptically removed, where half was retained for histopathology and half was collected for the immune assay (see paragraph 3.8).

The brain, heart, kidneys, liver, ovaries, uterus, spleen, and thymus were stored in 10% buffered formalin for fixation. The testes and right epididymis from each animal were placed in Davidson's fixative overnight (no longer than 24 hours). After fixation, the tissues were rinsed with deionized water and stored in 70% ethanol.

3.7 Sperm Analysis

Sperm analysis was only conducted for P generation animals. The left epididymis from each male was removed during necropsy and submitted for sperm analysis. In preparation, 10 mL of Gibco Medium 199 (M199) was pipetted into each well of Corning Costar 6-well cell culture plates and warmed on a slide warmer to approximately 37°C. Each epididymis was weighed, placed in a well containing M199, minced using small scissors, and incubated for 5 minutes at approximately 37°C. Then, samples were loaded on a standard count chamber slide (Leja). Slides were immediately loaded into a Hamilton-Thorne IVOS Sperm Analysis System and read using the IDENT program.

3.8 Immune Assay

Immune function was evaluated via a modified Jerne plaque forming assay (Cunningham 1976). Briefly, at necropsy, spleens were collected aseptically and prepared into single cell suspensions. The cell count of each spleen was determined using an automated cell counter (Bio-Rad TC20) and samples were diluted to approximately 2-E6 cell/mL in cold RPMI-1640 without Fetal Bovine Serum. The cell suspensions were combined with RPMI-1640 at 1:5, 2:5, and 5:5 for control/low, medium, and high dose groups, respectively. SRBCs (25%, 100 µL) and guinea pig complement (Cedarlane, Burlington, North Carolina 50 µL) were added to the cell suspension/RPMI-1640. The mixture was vortexed, pipetted onto parafilm, and loaded into two chambers of pre-made counting slides via capillary action. Slides were sealed with wax and incubated at 37°C for 2 hours. Plaques were counted under low power (10X) using a simple light microscope. Plaque forming cells were determined as plaque forming cells per million cells.

3.9 Histopathology

Preserved tissues were processed using standard techniques, sectioned 4–5 µm thick, and stained with hematoxylin and eosin using a routine automatic stainer. Testis and epididymis were additionally stained with PAS stain. Histopathology was performed for all high-dose and control animals at the discretion of the pathologist and based on observed toxicity and gross pathology findings. Lower dose groups were examined if exposure-related effects were seen in the high-dose group, gross lesions were present, or other signs of organ toxicity were noted (e.g., changes in organ mass). Histopathologic findings were subjectively graded across the following 6-point scale:

- Grade 0 (essentially normal) referred to tissues with changes observed in <1% of the sampled tissue.

- Grade 1 (minimum) referred to a change which affected <5% of the presented tissue area.
- Grade 2 (mild) referred to a change which affected 6 to 20% of the tissue area.
- Grade 3 (moderate) was scaled to refer to a change which affected 21 to 40% of the tissue area.
- Grade 4 (marked) was scaled for lesions affecting 41 to 80% of the tissue area.
- Grade 5 (severe) indicated that >80% of the tissue was affected (Schafer et al. 2018).

3.10 Hormone Analysis

Estradiol, progesterone, testosterone, triiodothyronine (T3), and total thyroxine (T4) were determined using the MAGPIX[®] analyzer and the Milliplex[®] Multi-Species Hormone Magnetic Bead Panel. The assay was conducted according to the manufacturer's instructions (Millipore 2013), and quality control standards were within the target reference ranges.

3.11 Data Collection and Statistical Analyses

Experimental data generated during the course of this study were recorded by hand and tabulated, summarized, and/or statistically analyzed using Microsoft Excel, Minitab, or SPSS. Environmental data were automatically recorded using MetaSys Building Management System.

Data not normally distributed were log transformed and retested for normality. Variance equality was determined by Levene's test. Analyses were conducted for males and females separately. Parameters measured multiple times (i.e., body weight) were analyzed using repeated measures one-way ANOVA, and those measured at the end of the study (i.e., organ weights) were analyzed using ANOVA with dose group as the main effect and age as a covariate. Absolute organ mass was analyzed by ANCOVA using dose group as the main effect and body weight at necropsy as the covariate (Bailey et al. 2004). Fisher's exact test was used to determine significant differences between treated and control groups for nominal or count data (e.g., histology). Statistical significance are defined at the $p < 0.05$ level. If dose effect was significant ($p < 0.05$), appropriate post hoc analyses were performed (Tukey's multiple comparison (ANOVA) or Sidak (for ANCOVA). Interpretation of changes in absolute organ mass, organ-to-body mass ratio, and organ-to-brain mass ratio in the evaluation of compound-related effects was based on published analysis of control animal data (Bailey et al. 2004).

3.12 Derivation of Toxicity Reference Values (TRVs)

TRVs will be derived according to standard procedure (APHC 2019). Briefly, either the Benchmark dose approach or the No Observed Adverse Effect Level (NOAEL) and Lowest Observed Adverse Effect Level (LOAEL) NOAEL/LOAEL approach will be used. Data that have at least three exposure groups plus a control and show a clear dose-response relationship using a unimodal design is best used to derive two TRVs, which is based on the benchmark dose approach. Data that lack the required number of dose groups or do not have a clear dose response relationship within well-designed and conducted parameters should be used to derive two TRVs, which one is based on the NOAEL and the other is based on the LOAEL. Where data are scarce and cannot be used for the aforementioned procedures, then the second

Toxicological Study No. S.0043781-18 and S.0059303-19, October 2017–August 2019

approach will be approximation with the use of uncertainty factors (UFs) to derive TRVs that estimate an NOAEL and/or LOAEL.

APPENDIX A

REFERENCES

- APHC (Environmental Health Risk Assessment Program). 2019. Technical Guide 254 (APHC TG 254, 1st revision). Standard Practice for Wildlife Toxicity Reference Values. Prepared by: Aberdeen Proving Ground, Maryland,
- Bailey, SA, RH Zidell, and RW Perry. 2004. Relationships between organ weight and body/brain weight in the rat: what is the best analytical endpoint? *Toxicol Pathol* 32 (4):448-66.
- CFR. 2011. Toxic Substances and Control Act (TSCA) 792 - Good Laboratory Practice Standards. In *Title 40*,: Code of Federal Regulations (CFR).
- Crossland, JP. 2006. Peromyscus - A Fascinating Laboratory Animal Model. In *AALAS Tech Talk*: AALAS.
- Cunningham, AJ. 1976. Evolution in microcosm: the rapid somatic diversification of lymphocytes. *Ann Immunol (Paris)* 127 (3-4):531-49.
- Hays, LE, CD Carpenter, and SL Petersen. 2002. Evidence that GABAergic neurons in the preoptic area of the rat brain are targets of 2,3,7,8-tetrachlorodibenzo-p-dioxin during development. *Environ Health Perspect* 110(Suppl 3):369-76.
- McCaffrey, KA, B Jones, N Mabrey, B Weiss, SH Swan, and HB Patisaul. 2013. Sex specific impact of perinatal bisphenol A (BPA) exposure over a range of orally administered doses on rat hypothalamic sexual differentiation. *Neurotoxicology* 36:55-62.
- Millipore. 2013. Multi-Species Hormone Magnetic Bead Panel. Billerica, MA: EMD Millipore Corp.
- NRC. 2011. *Guide for the Care and Use of Laboratory Animals*. 8th ed. Washington, D.C.: National Academies Press.
- Patisaul, HB, KT Burke, RE Hinkle, HB Adewale, and D Shea. 2009. Systemic administration of diarylpropionitrile (DPN) or phytoestrogens does not affect anxiety-related behaviors in gonadally intact male rats. *Horm Behav* 55(2):319-28.
- Plumel, MI, A Stier, D Thierse, A van Dorsselaer, F Criscuolo, F Bertile. 2014. Litter size manipulation in laboratory mice: an example of how proteomic analysis can uncover new mechanisms underlying the cost of reproduction. *Frontiers in Zoology* 11(41).
- Reiner, JL, SF Nakayama, AD Delinsky, JP Stanko, SE Fenton, AB Lindstrom, and MJ Strynar. 2009. Analysis of PFOA in dosed CD1 mice. Part 1. Methods development for the

Toxicological Study No. S.0043781-18 and S.0059303-19, October 2017–August 2019

analysis of tissues and fluids from pregnant and lactating mice and their pups. *Reprod Toxicol* 27(3-4):360-364.

Schafer, KA, J Eighmy, JD Fikes, WG Halpern, RR Hukkanen, GG Long, EK Meseck, DJ Patrick, MS Thibodeaux, CE Wood, and S Francke. 2018. Use of Severity Grades to Characterize Histopathologic Changes. *Toxicol Pathol* 46(3):256-265.

Schmidt, CM and WR Hood. 2012. Calcium Availability Influences Litter Size and Sex Ratio in White-Footed Mice (*Peromyscus leucopus*). *PLoS One* 7(8):e41402.

Sundstrom, M, SC Chang, PE Noker, GS Gorman, JA Hart, DJ Ehresman, A Bergman, and JL Butenhoff. 2012. Comparative pharmacokinetics of perfluorohexanesulfonate (PFHxS) in rats, mice, and monkeys. *Reprod Toxicol* 33(4):441-451.