Comments Regarding the Systematic Review of Immunotoxicity Associated with Perfluorooctanoic Acid (PFOA) and Perfluorooctane Sulfonate (PFOS)

Prepared on behalf of
3M
I-94 and McKnight Road
St. Paul, MN 55144-1000

July 5, 2016

GRADIENT

www.gradientcorp.com
20 University Road
Cambridge, MA 02138
617-395-5000

Table of Contents

Abbreviations

AhR	Airway Hyperresponsiveness
AMSTAR	A Measurement Tool to Assess Systematic Reviews
IgG	Immunoglobulin G
IgM	Immunoglobulin M
MOA/HR	Mode-of-Action Human Relevance
NTP	National Toxicology Program
OHAT	Office of Health Assessment and Translation
PFOA	Perfluorooctanoic Acid
PFOS	Perfluorooctane Sulfonate
RoB	Risk of Bias
SRBC	Sheep Red Blood Cells

Executive Summary -

These comments on the Systematic Review of Immunotoxicity Associated with Perfluorooctanoic Acid (PFOA) and Perfluorooctane Sulfonate (PFOS) (NTP, 2016) have been prepared on behalf of 3M by Gradient, an environmental consulting firm with particular expertise in toxicology and risk assessment. We note that, while this work has been sponsored by 3M, the conclusions are solely those of the authors. Because of the limited time to comment on the document, rather than conducting a broad analysis, we have focused on the toxicology studies with discussion of human studies to illustrate concerns with application of methodology, integration of evidence and conclusions.

Our overall conclusion is that, while the review is comprehensive in its coverage of the immunotoxicology and epidemiology of PFOA and PFOS, there are a number of issues with the National Toxicology Program's (NTP's) methodology and conclusions. Much of the human and animal evidence does not support NTP's conclusions, and the hazard ratings for both PFOA and PFOS should be downgraded. We present a brief summary of the basis for this overall conclusion below, with detailed support later in this document.

The Office of Health Assessment and Translation (OHAT) framework is insufficient to produce a balanced systematic review.

- The OHAT framework should require the consideration of mechanistic data in reaching overall conclusions for the evidence of health effects.
- The Risk of Bias (RoB) criteria are not given appropriate weight in evidence integration.
- The OHAT framework does not provide explicit guidance with regard to evidence integration.

NTP does not appear to follow OHAT guidance.

- NTP does not adequately support its high confidence rating for the level of evidence for an association between PFOA and PFOS and antibody response and hypersensitivity.
- NTP does not categorize studies into tiers of quality, causing its interpretation of the confidence in the realms of evidence (human, animal, *in vitro*) to be unclear.

NTP's analysis of the human evidence for PFOA and PFOS immune effects is biased, with selective presentation of results and a focus on positive over null results.

- NTP emphasizes statistically significant positive study results over null results, leading to biased interpretations of the body of evidence.
- NTP does not adequately address the lack of dose-response relationships and inconsistent results both within and across studies.

NTP's analysis of the animal evidence for PFOA and PFOS immunotoxicity is biased, with selective presentation of results and a focus on positive over null results.

- The high risk of bias NTP has identified in the animal studies for antibody response should not be offset by the evidence of dose-response, which, in most studies, is present only in association with systemic toxicity. The confidence rating for antibody response for both PFOA and PFOS should be downgraded based on the high risk of bias.
- NTP has determined a high confidence rating for hypersensitivity effects of PFOA based on two animal studies that had contradictory findings. One of those studies tested only one PFOA dose, and the other did not result in a typical dose-response and had a high risk of bias. The confidence rating for hypersensitivity should be downgraded.

The integration of the evidence is not conducted in a transparent manner, and is biased, resulting in overly high levels of confidence ascribed to the evidence.

- NTP does not provide a sufficiently detailed, transparent description of the process by which evidence was integrated to reach the conclusion that PFOA and PFOS are presumed to be immune hazards. There is no discussion that integrates the weight of evidence for all immune endpoints examined. Further, there is no specific "Evidence Integration" section for compiling and comparing the human, animal, *in vitro*, and other data across all endpoints.
- NTP emphasizes the positive results over null results, leading to a biased analysis that overstates the strength of evidence.
- The RoB criteria are not given appropriate weight in evidence integration. The main body of evidence on which NTP relies (animal data) has a serious risk of bias, which may have led to a distorted relationship between exposure and disease.
- Based on the lack of evidence for decreased disease resistance in humans and limited evidence in animals, the confidence rating of "presumed to be an immune hazard" for immunosuppression of PFOA and of PFOS has inadequate support and should be downgraded.

1 Summary of NTP's Conclusions Regarding the Immunotoxicity of PFOA and PFOS

1.1 Perfluorooctanoic Acid (PFOA)

The National Toxicology Program (NTP) (2016) concludes that perfluorooctanoic acid (PFOA) is presumed to be an immune hazard to humans based on evidence for the suppression of the antibody response. This conclusion is derived from a moderate level of confidence in the body of evidence from four prospective and two cross-sectional human studies, and a high level of confidence in the body of evidence from seven studies in mammals. Relevant mechanistic data, including *in vitro* studies, are not considered to provide evidence to support or refute the biological plausibility of the association of PFOA with this endpoint. Based on a moderate level of confidence in the human evidence and a high level of confidence in the animal evidence, NTP (2016) concludes that PFOA is presumed to be an immune hazard to humans.

NTP (2016) also concludes that PFOA is presumed to be an immune hazard to humans based on evidence for increased hypersensitivity. This conclusion is derived from a low level of confidence in the body of evidence from five cross-sectional child exposure studies, and a high level of confidence in the body of evidence from two studies in mammals. Relevant mechanistic data, including *in vitro* studies, are not considered to provide evidence to support or refute the biological plausibility of the association of PFOA with this endpoint. Based on a low level of confidence in the human evidence and a high level of confidence in the animal evidence, NTP (2016) concludes that PFOA is presumed to be an immune hazard to humans.

NTP (2016) has not found sufficient evidence to make a determination as to whether or not PFOA was an immune hazard based on autoimmune effects.

1.2 Perfluorooctane Sulfonate (PFOS)

NTP (2016) concludes that PFOS is presumed to be an immune hazard to humans based on evidence for the suppression of the antibody response. This conclusion is derived from a moderate level of confidence in the body of evidence from four prospective and two cross-sectional human studies, and a high level of confidence in the body of evidence from eight studies in mammals. Relevant mechanistic data, including *in vitro* studies, were not considered to provide evidence to support or refute the biological plausibility of the association of PFOS with this endpoint. Based on a moderate level of confidence in the human evidence and a high level of confidence in the animal evidence, NTP (2016) concludes that PFOS is presumed to be an immune hazard to humans.

NTP (2016) has not found sufficient evidence to make a determination as to whether or not PFOS is an immune hazard based on hypersensitivity or autoimmune effects.

2 General Comments

NTP (2016) conducted a thorough review of the immunotoxicity literature concerning PFOA and PFOS. Its descriptions of the individual studies are generally accurate, and it appropriately considers primary *vs*. secondary immune effects associated with these perfluorinated chemicals. The literature search and review are consistent with the framework prescribed by NTP's Office of Health Assessment and Translation (OHAT; NTP, 2015). However, the integration of the body of evidence is not conducted in a transparent manner, and is biased, resulting in overly high levels of confidence ascribed to the evidence.

2.1 Methodology

2.1.1 Comments on the OHAT Framework

Overall, specific aspects of the OHAT framework are thorough, prescriptive (*i.e.*, based on a clearly specified method), and appropriate as written; however, other areas, most notably the methods for evidence integration, are insufficient to produce a balanced systematic review.

First, the OHAT framework should require the consideration of mechanistic data in reaching overall conclusions for the evidence of health effects. Currently, the framework notes that incorporation of mechanistic data is "optional." In addition, when mechanistic data are used, the data are only considered after NTP reaches an initial conclusion on the hazard rating ("known," "suspected," "presumed," "not classifiable," or "not identified as a hazard") based on human and animal evidence. The mechanistic data may then be used to up- or downgrade the classification. In a balanced systematic review, mechanistic data should be treated the same as human and animal data; in other words, they should be assessed for quality and relevance to determine whether they support a biologically plausible association between the agent and health effect of interest. There are available frameworks for incorporating mechanistic data in hazard and risk assessment (see, for example, the Mode-of-Action Human Relevance ["MOA/HR"] Framework, Meek *et al*., 2014).

In addition, while the RoB criteria are well thought out and specific, the results are not given appropriate weight in evidence integration. NTP developed a very detailed system for assessing RoB in their systematic reviews, and the guidance appears to place much emphasis on conducting a thorough RoB analysis within a review. However, the OHAT guidance subsequently fails to describe how the RoB analysis should be incorporated into the integration of evidence to reach a final hazard classification. This is a serious limitation in that the lack of requirement for its incorporation allows for the RoB conclusions to be discarded in the final hazard identification. Numerous policymaking and guidance agencies, including the National Research Council, consider study quality assessment a critical step in the systematic review process (Lynch *et al*., 2015; Stephens *et al*., 2016). Further, these systems specifically state that the results of a study quality analysis should be used in formulating conclusions about hazard or risk potential. Several systems used to assess the quality of systematic reviews have also highlighted the importance of study quality in hazard and risk assessment. For example, one of the eleven key questions of the AMSTAR system ("A Measurement Tool to Assess Systematic Reviews," developed by an interdisciplinary team of research professionals from Canada and the EU) is, "Was the scientific quality of the included studies used appropriately in formulating conclusions?" (Shea *et al*., 2007). NTP should

strengthen its OHAT framework with more prescriptive requirements for the consideration of study quality in hazard identification determinations.

2.1.2 Application of the OHAT Framework in the Systematic Review of PFOA and PFOS

Overall, NTP's (2016) systematic review of PFOA and PFOS is generally consistent with the OHAT framework. However, there are some important exceptions; most notably, NTP (2016) does not appear to follow guidance with regard to the confidence rating of the level of evidence. The NTP guidance (NTP, 2015) states that "high confidence" means NTP has high confidence in the association between exposure to the substance and the outcome, and specifically that "the true effect is highly likely to be reflected in the apparent relationship [emphasis in original]." Further, the guidance specifies that a high confidence rating is reserved for instances in which further research is very unlikely to change confidence in the apparent relationship. As detailed in Sections 3 and 4 below, NTP (2016) does not adequately support its high confidence rating for the level of evidence for an association between PFOA or PFOS and antibody response and hypersensitivity. The main body of evidence on which it relies (animal data) has a serious risk of bias, and by definition, bias (*e.g.*, poor outcome assessment) is an error that may lead to a distorted relationship between exposure and disease (Rothman *et al.*, 2008). In other words, bias reduces confidence that the true effect is reflected in the study results.

In addition, regarding the RoB analysis in the current assessment, NTP (2016) does not categorize studies into tiers of quality. The OHAT framework recommends a tiering approach with 3 tiers based on overall study quality; these tiers are used to determine whether the overall RoB in each realm of evidence is "not likely," "serious," or "very serious." These tiers are then used to facilitate interpretation of the confidence in the realms of evidence (human, animal, *in vitro*). NTP (2016) determined, for example, that the bodies of animal evidence for antibody response after exposure to PFOS and PFOA had a "serious concern" for RoB. The OHAT guidance states that this classification should be assigned in situations where there is "plausible bias that raises some doubt about the results," and "most information is from Tier 1 and 2 studies" (NTP, 2015), yet NTP (2016) does not assign tiers to the studies in the PFOA and PFOS assessments.

Finally, the NTP (2016) review of PFOA and PFOS immunotoxicity fails to fully describe the potential for publication bias in the available literature. Publication bias results when published research does not reflect the overall population of completed research studies (published and unpublished). Typically, this occurs because journal editors are more likely to publish positive findings than null findings (*i.e.*, those reporting no associations). As a result, studies reporting statistically significant, positive findings are generally overrepresented in the body of literature.

The OHAT framework includes publication bias as one of the key aspects in considering confidence in the research literature, and the OHAT guidance document states that suspected publication bias may be grounds for downgrading confidence in a body of evidence. While it may be difficult to determine whether publication bias has occurred, there are some indicators of the potential for this bias. As noted in the NTP guidance, one possible indicator is abstracts or other types of grey literature identified in the literature search process that do not appear as full-length articles "within a reasonable time frame (around 3 to 4 years)." The presence of few, small-sized studies published in recent years may also suggest possible publication bias, due to a publication lag for studies showing no effects (NTP, 2015). In the review of PFOA and PFOS, NTP (2016) indicates in its evidence tables that publication bias neither increased nor decreased confidence in any endpoint; however, the text provides no indication that NTP performed an analysis and if it did, what the results showed. NTP (2016) should address this important issue in its analysis of the potential immunotoxic effects of PFOS and PFOA.

2.2 Other Comments on Methodology

Review articles can be an important source of information for the integration of evidence. While it is important for an agency to conduct a review of the literature without the influence of review authors' opinions, during the integration of evidence stage, reviews can be informative as to alternate hypotheses that may be derived from a body of evidence and should be considered as part of the integration process. For example, NTP (2016) cites the DeWitt *et al*. (2012) review of animal immunotoxicity of perfluorinated compounds several times throughout the Systematic Review. There is at least one other recent and relevant review of human studies, by Chang *et al*. (2016), available as of January 2016, that is not cited by NTP (2016). While the DeWitt *et al*. (2012) conclusions were consistent with many of NTP's (2016) conclusions, the Chang *et al*. (2012) review concluded that the human evidence was insufficient to reach a conclusion about a causal relationship between exposure to PFOA and PFOS and any immune related health condition in humans. The omission of the Chang *et al*. (2012) review suggests a potential bias in the literature cited by NTP (2016).

3 Analysis of PFOA

NTP (2016) concludes that PFOA is *presumed to be an immune hazard to humans* based on evidence for both immunosuppression and hypersensitivity. Overall, the interpretation of the body of human, animal, and *in vitro* evidence is biased, with selective presentation of results and a focus on positive over null results.

3.1 Human Evidence

The overall description of the human evidence provided by NTP (2016) on PFOA and immunological outcomes are generally accurate and thorough. However, NTP (2016) emphasizes statistically significant positive study results over null results, leading to biased interpretations of the body of evidence. In the data tables within the body of the report, NTP (2016) presents selected positive results for each outcome assessed, while null results are relegated to a separate section at the end of the document. For instance, in Table 11, NTP (2016) presents only the Looker *et al*. (2014) results for the PFOA metrics that demonstrated a significant reduction in antibody response, and omitted several metrics with non-significant outcomes. Additionally, results are presented for the significant findings with respect to Influenza A/H3N2 vaccine, while the nonsignificant results for Influenza A/H1N1 and Influenza B vaccines are only available in a separate data section at the end of the report. NTP (2016) does not report anywhere in the document the odds of seroconversion for any of the vaccine components, in which no associations were observed with increasing quartiles of PFOA. There is also no discussion with respect to the lack of a clear dose-response relationship in the measures with statistically significant results in the Looker *et al*. (2014) study, nor with respect to the likelihood of false positive results due to the multiple comparisons included in this study. Similarly, in Table 16, only the PFOA metric that produced statistically significant results in the Humblet *et al*. (2014) study is presented while the nonsignificant results are presented in a separate data section at the end of the report. Focusing attention on positive results and failing to discuss or present nonsignificant findings in the main report results in a biased representation of the body of evidence. Overall, considering the inconsistent results reported for humans exposed to PFOA, we conclude that NTP's (2016) confidence ratings of moderate for antibody response, low for infectious disease, and low for hypersensitivity are appropriate.

3.2 Animal Evidence

3.2.1 Immunosuppression

NTP's (2016) conclusions regarding immunosuppression specifically refer to reduced IgM production in the sheep red blood cell (SRBC) assay, which is widely accepted as a standard immunotoxicological evaluation and is considered a primary immunological endpoint by NTP. NTP (2016) identifies seven key studies as providing evidence of immunosuppression and assigns a high level of initial confidence to the evidence of reduced antibody production (DeWitt *et al*., 2008, 2009, 2016; Hu *et al*., 2010; Loveless *et al*., 2008; Vetvicka and Vetvickova, 2013; Yang *et al*., 2002). NTP (2016) identifies a number of methodological issues in these studies, and concludes that a high RoB is probable in the studies, thus reducing the confidence in the evidence. However, NTP (2016) considers this confidence reduction to be offset by clear evidence of a dose-response, leading to a final confidence rating of "High."

Only one of the seven key studies provides clear evidence of a dose-response decrease in SRBC-specific Immunoglobulin M (IgM) production in the absence of systemic toxicity in all animals (DeWitt *et al*., 2008). For this reason, we conclude that the evidence for a dose-response is not sufficient to offset the high RoB that was identified by NTP (2016) as "probably" present in all of the key studies (Figure D13). As no other factor was identified to increase the confidence in the evidence, the final level of confidence in the evidence that PFOA causes immunosuppression should be downgraded. This proposal to downgrade is supported by the limitations of the following studies cited by NTP (2016):

- DeWitt *et al*. (2008): As noted by the authors, the experimental serum PFOA levels associated with a significant antibody reduction were 15,000 times the serum levels of the general population.
- DeWitt *et al*. (2009): SRBC-specific IgM antibody titers were not significantly different from controls following exposure to PFOA up to 15 mg/kg in drinking water for 10 days.
- DeWitt *et al*. (2016): SRBC-specific IgM antibody titers in both wild-type and peroxisome proliferator-activated receptor α knockout mice did not differ from control values following exposure to PFOA up to 30 mg/kg in drinking water for 15 days.
- Hu *et al*. (2010): No statistical differences were observed in SRBC-specific IgM antibody titers in female offspring following dam exposure up to 1 mg/kg PFOA from gestation days 6 to 17.
- Loveless *et al*. (2008): SRBC-specific IgM antibody titers were not significantly different from controls following exposure to PFOA up to 30 mg/kg by oral gavage in rats for 29 days. Mice were found to have decreased SRBC-specific IgM antibody titers following PFOA exposures of 10 and 30 mg/kg, however these doses were also associated with systemic toxicity as evidenced by reduced body weight and elevated markers of stress.
- Vetvicka and Vetvickova (2013): Only a single dose level of PFOA (20 mg/kg-day by oral gavage) was tested, and therefore dose-response cannot be evaluated.
- Yang *et al*. (2002): Only a single dose level of PFOA (~40 mg/kg-day in diet) was tested, and therefore dose-response cannot be evaluated.

Overall, the studies that show an association of PFOA with reduced antibody production have high RoB and provide very limited evidence for a dose-response effect in the absence of systemic toxicity. For these reasons, the level of confidence in the animal evidence for immunosuppression effects of PFOA should be downgraded.

3.2.2 Hypersensitivity

NTP (2016) cites two animal studies to provide evidence for a rating of high confidence that PFOA can be presumed to cause hypersensitivity (Fairley *et al*., 2007; Ryu *et al*., 2014).[1] The studies by Ryu *et al*. (2014) and Fairley *et al*. (2007) evaluated airway hyperresponsiveness (AhR) in mice, among other endpoints. The findings of these studies are contradictory and do not provide clear evidence of a dose-response relationship. Furthermore, the study by Fairley *et al*. (2007) has a high RoB, while the study by Ryu *et al*. (2014) only observed effects at high serum levels.

While both studies reported a statistically significant increase in AhR in mice exposed to PFOA, the types of AhR observed were different between studies. Ryu *et al*. (2014) reported an unusual (macrophage-

[1] NTP also cites Singh *et al*. (2012) in the main text as supporting evidence but does not list the study in Tables 6 or 17, nor does it evaluate the study for RoB. Thus, it is unclear how NTP used the study in their evaluation.

mediated) non-allergic increase in AhR in mice exposed to PFOA during and after gestation, but did not observe an enhanced allergic hypersensitivity. In contrast, Fairley *et al*. (2007) observed an enhancement of allergen-induced airway hyperresponsiveness but no effect of PFOA exposure alone.

The data lack a clear dose-response for immune hypersensitivity. Ryu *et al*. (2014) used only a single dose. Although the study by Fairley *et al*. (2007) used a range of doses, an unexpected inverted-U-shaped dose-response was observed. The study by Fairley *et al*. (2007) also has several significant methodological issues that increase the RoB. NTP (2016) identified many of these issues (see Figure D36) but failed to note that animals were "assigned to homogenous weight groups prior to the first exposure" and not randomly assigned. Lastly, the experimental serum PFOA levels associated with statistically significant effects in the study by Ryu *et al*. (2014) were 1,000 times those observed in the general human population (Calafat *et al*., 2007).[2]

Given the unexplained inconsistencies in the effects of PFOA exposure, high exposure levels, questionable dose response evidence, and a high RoB in one of the two studies, the confidence level for PFOA effects on hypersensitivity should be downgraded.

3.3 Integration of Evidence

NTP (2016) concludes that exposure to PFOA is *presumed to be an immune hazard to humans* based on a moderate level of evidence in humans and a high level of evidence in animals for immunosuppression, and a low level of evidence in humans and high level of evidence in animals for hypersensitivity. However, NTP (2016) does not provide a sufficiently detailed, transparent description of the process by which it integrates evidence to reach the conclusion that PFOA is presumed to be an immune hazard. The "Evidence Synthesis" sections presented at the end of each endpoint are approximately two to three paragraphs, and there is no discussion at the end of the larger PFOA health effects section that integrates the weight of evidence for all immune endpoints examined. Further, there is no specific "Evidence Integration" section for considering all together the human, animal, *in vitro*, and other data across all endpoints. The weight of evidence from all realms of evidence are discussed only briefly in the document's conclusion section. In the absence of a sufficient evidence integration discussion, NTP (2016) does not adequately support its final hazard identification conclusions.

With regard to antibody response, NTP (2016) concludes that there is moderate confidence that exposure to PFOA is associated with suppression of the antibody response in humans, "based on consistent suppression in at least one measure of the anti-vaccine antibody response across multiple studies with evidence from prenatal, childhood, and adult exposures to PFOA." However, NTP (2016) emphasizes the positive results over the null results, leading to a biased analysis that overstates the strength of evidence. With regard to animal data, despite a high RoB across studies, NTP (2016) concludes that it has high confidence in the evidence. NTP (2016) minimizes the issues of RoB with the assertion that there is a strong dose-response relationship between PFOA and antibody response in animals; however, only one of seven key studies showed such a relationship in the absence of systemic toxicity. Finally, the available mechanistic data were insufficient to provide evidence to support that PFOA suppresses antibody response.

With respect to PFOA exposure and hypersensitivity-related outcomes, NTP (2016) overstates the confidence in the animal evidence and the body of evidence as a whole. NTP (2016) assigns a low confidence rating to the human evidence of hypersensitivity based on three cross-sectional studies, and a high confidence rating to the animal evidence based on only two studies, despite important design flaws

[2] Serum PFOA levels were not quantified in the study by Fairley *et al*. (2007).

highlighted in the RoB analysis (*e.g.,* exposure characterization and outcome assessment) and unexplained inconsistencies in the results. NTP (2016) does not sufficiently support the conclusion that, despite inconsistencies in the results, and a high RoB in the only study with more than one dose level, it has high confidence in the animal data for that endpoint. Finally, NTP (2016) concludes the mechanistic (*i.e.*, *in vitro*) data are not adequate to support or refute the biological plausibility of the endpoint, so these data could not be used to increase confidence in the evidence for hypersensitivity.

Across endpoints, given the weak human and mechanistic data and lower quality, limited, and inconsistent animal data, the evidence is insufficient to conclude that PFOA is an immune hazard in humans, and it is unclear how NTP (2016) reached its final hazard identification conclusions.

4 Analysis of PFOS

NTP (2016) concludes that PFOS is *presumed to be an immune hazard to humans* based on evidence for immunosuppression. Overall, the interpretation of the body of human, animal, and *in vitro* evidence is biased, with selective presentation of results and a focus on positive over null results.

4.1 Human Evidence

The overall descriptions of the human evidence provided by NTP (2016) on PFOS and immunological outcomes are generally accurate and thorough. However, NTP (2016) emphasizes statistically significant positive study results over null results, leading to biased interpretations of the body of evidence. In the data tables within the body of the report, NTP (2016) presents selected positive results for each outcome assessed, while null results are relegated to tables in a separate data section at the end of the report. For instance, in Table 21, NTP (2016) presents only the Fei *et al*. (2010) results for the PFOS metric stratified by sex, which showed a significant increase in hospitalization for infectious disease. The age-stratified results, which showed a lower risk of hospitalization for infections during the first year of life and no dose response pattern, are available only in a separate data section at the end of the report. Focusing attention on positive results and failing to discuss or present nonsignificant findings in the main report results in a biased representation of the body of evidence. Overall, considering the inconsistent results reported for humans exposed to PFOS, the confidence rating of moderate for immunosuppression, low for infectious disease, and very low for hypersensitivity is appropriate.

4.2 Animal Evidence

NTP's (2016) conclusions on immunosuppression are primarily based on reduced IgM production in the SRBC assay, which is widely accepted as a standard immunotoxicological evaluation and is considered a primary immunological endpoint by NTP. NTP (2016) identifies eight key studies as providing evidence of immunosuppression and assigns a high level of initial confidence to the evidence of reduced antibody production (Dong *et al*., 2009, 2011; Keil *et al*., 2008; Lefebvre *et al*., 2008; Peden-Adams *et al*., 2008; Qazi *et al*., 2010; Vetvicka and Vetvickova, 2013; Zheng *et al*., 2009). NTP (2016) identifies a number of methodological issues in these studies that may increase the likelihood of observing an effect (*e.g.*, a biased experimental design), and concludes that a high RoB is probable in the studies, thereby reducing the confidence in the evidence. However, NTP (2016) considers this confidence reduction to be offset by clear evidence of a dose-response, leading to a final confidence rating of "High."

A dose-dependent reduction in SRBC-specific IgM antibody titers was observed in studies by Dong *et al*. (2009, 2011), Peden-Adams *et al*. (2008), and Zheng *et al*. (2009). However, these effects paralleled systemic toxicity in the studies by Dong *et al*. (2009, 2011) and Zheng *et al*. (2009). Thus, only the study by Peden-Adams *et al*. (2008) provides evidence of a dose-response effect for T-dependent antigen response. For this reason, we conclude that the evidence for a dose-response is not strong enough to offset the high RoB that is identified by NTP (2016) as "probably" present in all of the key studies (Figure D15). As no other factors were identified to increase the confidence in the evidence, the final level of confidence in the evidence that PFOS causes immunosuppression should be downgraded. This is supported by the limitations of the following studies cited by NTP (2016):

- Lefebvre *et al*. (2008): No differences in serum keyhole limpit hemocyanin-specific Immunoglobulin G (IgG) levels were observed in rats treated with PFOS up to 100 mg/kg-day in the diet.
- Qazi *et al*. (2010): No differences in SRBC-specific IgM levels were observed in mice treated with 7 mg/kg PFOS for 28 days in the diet.
- Vetvicka and Vetvickova (2013): Only a single dose level of PFOS (20 mg/kg-day by oral gavage) was tested, and therefore dose-response cannot be evaluated.

The study by Peden-Adams *et al*. (2008) presents compelling data that PFOS may cause immunotoxicity at experimental serum levels within the range of the general human population. However, these results have not been replicated in publications by this or other research groups. In subsequent publications co-authored by Dr. Peden-Adams (Mollenhauer *et al*., 2011; Fair *et al*., 2011), minimal effects to the immune system were observed in association with PFOS treatment. Neither study attempted to reproduce the results of the original Peden-Adams *et al*. (2008) publication, despite the original study consisting of only a single experiment of five mice per gender per group. This is discussed by Fair *et al*. (2011):

> Moreover, data from this study confirm that numbers of CD4+ cells were within normal ranges. This contrasts a previous report from this laboratory where absolute numbers of CD4+ cells were decreased in female B6C3F1 mice at 0.1 mg/kg [total administered dose] but not at 1.0 mg/kg [total administered dose] using a similar 28-day exposure regimen (Peden-Adams et al., 2008). This previous observation was from a single experiment, whereas in this study, the experiment was repeated twice for absolute numbers and three times for percent changes with all experiments yielding the same results. The effect previously reported was not dose-responsive and is likely to be a transient effect. Overall, these data indicate that T-helper cells, B-cells, and MHC-II+ cells were not selectively eliminated. (pp. 25-26)

Overall, the studies that show an association of PFOS with reduced antibody production have a high risk of bias, and there is very limited evidence for a dose-response effect in the absence of systemic toxicity. For these reasons, the level of confidence in the animal evidence for immunosuppression effects of PFOA should be downgraded.

4.3 Integration of Evidence

NTP (2016) concludes that "exposure to PFOS is *presumed to be an immune hazard to humans* based on a high level of evidence that PFOS suppressed the antibody response from animal studies and a moderate level of evidence from studies in humans," and notes that though weak, there is additional evidence that PFOS affects other aspects of the immune system. NTP (2016) does not, however, provide a sufficiently detailed, transparent description of the process by which the agency integrated evidence to reach this conclusion. The "Evidence Synthesis" sections presented at the end of each endpoint are approximately two to three paragraphs long, with very little discussion regarding the method by which the confidence in each body of evidence was brought together to make hazard identification conclusions. There is no discussion at the end of the PFOS health effects section that integrates the weight of evidence for all immune endpoints examined, and no specific "Evidence Integration" section that describes how the human, animal, *in vitro*, and other data were integrated. The weight of evidence from all realms of evidence is discussed only briefly in the overall conclusion section. NTP (2016) does not adequately support its final hazard identification conclusion that PFOS is presumed to be an immune hazard.

Overall, the available human data on PFOS and immunotoxicity are limited, and there are important uncertainties regarding the specificity of the immune effects in animals. As in its assessment of PFOA, NTP (2016) assigns a high confidence rating to the animal evidence even though the RoB analysis indicates that there were important limitations in all of the studies cited (*e.g.,* high RoB with regard to exposure characterization). In addition, the antibody response occurred in the presence of systemic toxicity in all but one study, which raises questions regarding whether PFOS caused a targeted immune response, or whether the immune toxicity was secondary to systemic toxicity. In addition, some of the reported results have not been reproduced, reducing confidence in the body of animal evidence. NTP (2016) does not sufficiently support the conclusion that one can have high confidence in the animal data despite the high RoB. With regard to human studies, NTP (2016) reports a moderate level of evidence of antibody response to PFOS in humans; however, it has low confidence in the human evidence of increased incidence of infectious disease. There are very few data on disease incidence after exposure to PFOS, and only two studies measured both antibody response and PFOS-associated increases in infectious disease. Therefore, even if PFOS exposure modifies antibody response in humans, there is currently insufficient evidence to conclude that this response will be associated with reduced ability to fight disease. Lastly, there is no established mechanism for PFOS-associated suppression of the antibody response, and the available mechanistic evidence are insufficient to support the biological plausibility of this effect. Overall, when weighing the human and animal data together, and considering the lack of an established mechanism, the evidence does not support that PFOS is an immune hazard, and the methodology that NTP (2016) uses to reach its final hazard identification conclusions is unclear.

References –

Calafat, AM; Kuklenyik, Z; Reidy, JA; Caudill, SP; Tully, JS; Needham, LL. 2007. "Serum concentrations of 11 perfluoroalkyl compounds in the U.S. population: Date from the National Health and Nutrition Examination Survey (NHANES) 1999-2000." *Environ. Sci. Technol.* 41(7):2237-2242.

Chang, ET; Adami, HO; Boffetta, P; Wedner, HJ; Mandel, JS. 2016. "A critical review of perfluorooctanoate and perfluorooctanesulfonate exposure and immunological health conditions in humans." *Crit. Rev. Toxicol.* 46(4):279-331. doi: 10.3109/10408444.2015.1122573.

DeWitt, JC; Copeland, CB; Strynar, MJ; Luebke, RW. 2008. "Perfluorooctanoic acid-induced immunomodulation in adult C57BL/6J or C57BL/6N female mice." *Environ. Health Perspect.* 116(5):644-650.

DeWitt, JC; Copeland, CB; Luebke, RW. 2009. "Suppression of humoral immunity by perfluorooctanoic acid is independent of elevated serum corticosterone concentration in mice." *Toxicol. Sci.* 109(1):106-112. doi: 10.1093/toxsci/kfp040.

DeWitt, JC; Peden-Adams, MM; Keller, JM; Germolec, DR. 2012. "Immunotoxicity of perfluorinated compounds: Recent developments." *Toxicol. Pathol.* 40(2):300-311. doi: 10.1177/0192623311428473.

DeWitt, JC; Williams, WC; Creech, NJ; Luebke, RW. 2016. "Suppression of antigen-specific antibody responses in mice exposed to perfluorooctanoic acid: Role of PPARa and T- and B-cell targeting." *J. Immunotoxicol.* 13(1):38-45. doi: 10.3109/1547691X.2014.996682.

Dong, GH; Zhang, YH; Zheng, L; Liu, W; Jin, YH; He, QC. 2009. "Chronic effects of perfluorooctanesulfonate exposure on immunotoxicity in adult male C57BL/6 mice." *Arch. Toxicol.* 83(9):805-815.

Dong, GH; Liu, MM; Wang, D; Zheng, L; Liang, ZF; Jin, YH. 2011. "Sub-chronic effect of perfluorooctanesulfonate (PFOS) on the balance of type 1 and type 2 cytokine in adult C57BL6 mice." *Arch. Toxicol.* 85(10):1235-1244.

Fair, PA; Driscoll, E; Mollenhauer, MA; Bradshaw, SG; Yun, SH; Kannan, K; Bossart, GD; Keil, DE; Peden-Adams, MM. 2011. "Effects of environmentally-relevant levels of perfluorooctane sulfonate on clinical parameters and immunological functions in B6C3F1 mice." *J. Immunotoxicol.* 8(1):17-29.

Fairley, KJ; Purdy, R; Kearns, S; Anderson, SE; Meade, BJ. 2007. "Exposure to the immunosuppressant, perfluorooctanoic acid, enhances the murine IgE and airway hyperreactivity response to ovalbumin." *Toxicol. Sci.* 97(2):375-383.

Fei, C; McLaughlin, JK; Lipworth, L; Olsen, J. 2010. "Prenatal exposure to PFOA and PFOS and risk of hospitalization for infectious diseases in early childhood." *Environ. Res.* 110(8):773-777. doi: 10.1016/j.envres.2010.08.004.

Hu, Q; Strynar, MJ; DeWitt, JC. 2010. "Are developmentally exposed C57BL/6 mice insensitive to suppression of TDAR by PFOA?" *J. Immunotoxicol.* 7(4):344-349.

Humblet, O; Diaz-Ramirez, LG; Balmes, JR; Pinney, SM; Hiatt, RA. 2014. "Perfluoroalkyl chemicals and asthma among children 12-19 years of age: NHANES (1999-2008)." *Environ. Health Perspect.* 122(10):1129-1133. doi: 10.1289/ehp.1306606.

Keil, DE; Mehlmann, T; Butterworth, L; Peden-Adams, MM. 2008. "Gestational exposure to perfluorooctane sulfonate (PFOS) suppresses immune function in B6C3F1 mice." *Toxicol. Sci.* 103(1):77-85.

Lefebvre, DE; Curran, I; Armstrong, C; Coady, L; Parenteau, M; Liston, V; Barker, M; Aziz, S; Rutherford, K; Bellon-Gagnon, P; Shenton, J; Mehta, R; Bondy, G. 2008. "Immunomodulatory effects of dietary potassium perfluorooctane sulfonate (PFOS) exposure in adult Sprague-Dawley rats." *J. Toxicol. Environ. Health A* 71(23):1516-1525.

Looker, C; Luster, MI; Calafat, AM; Johnson, VJ; Burleson, GR; Burleson, FG; Fletcher, T. 2014. "Influenza vaccine response in adults exposed to perfluorooctanoate and perfluorooctanesulfonate." *Toxicol. Sci.* 138(1):76-88. doi: 10.1093/toxsci/kft269.

Loveless, SE; Hoban, D; Sykes, G; Frame, SR; Everds, NE. 2008. "Evaluation of the immune system in rats and mice administered linear ammonium perfluorooctanoate." *Toxicol. Sci.* 105(1):86-96. doi: 10.1093/toxsci/kfn113.

Lynch, HN; Goodman, JE; Tabony, JA; Rhomberg, LR. 2015. "Systematic comparison of study quality criteria." *Regul. Toxicol. Pharmacol.* doi: 10.1016/j.yrtph.2015.12.017.

Meek, ME; Palermo, CM; Bachman, AN; North, CM; Lewis, RJ. 2014. "Mode of action human relevance (species concordance) framework: Evolution of the Bradford Hill considerations and comparative analysis of weight of evidence." *J. Appl. Toxicol.* doi: 10.1002/jat.2984.

Mollenhauer, MA; Bradshaw, SG; Fair, PA; McGuinn, WD; Peden-Adams, MM. 2011. "Effects of perfluorooctane sulfonate (PFOS) exposure on markers of inflammation in female B6C3F1 mice." *J. Environ. Sci. Health A* 46(2):97-101.

National Toxicology Program (NTP). 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), 98p., January 9. Accessed on June 27, 2016 at http://ntp.niehs.nih.gov/pubhealth/hat/noms/index-2.html.

National Toxicology Program (NTP). 2016. "Systematic Review of Immunotoxicity Associated with Exposure to Perfluorooctanoic Acid (PFOA) and Perfluorooctane Sulfonate (PFOS)." Office of Health Assessment and Translation (OHAT), 144p., June 6. Accessed on June 27, 2016 at http://ntp.niehs.nih.gov/ntp/about_ntp/monopeerrvw/2016/july/draftsystematicreviewimmunotoxicityassociatedpfoa_pfos_508.pdf.

Peden-Adams, MM; Keller, JM; EuDaly, JG; Berger, K; Gilkeson, GS; Keil, DE. 2008. "Suppression of humoral immunity in mice following exposure to perfluorooctane sulfonate (PFOS)." *Toxicol. Sci.* 104(1):144-154. doi: 10.1093/toxsci/kfn059.

Qazi, MR; Nelson, BD; Depierre, KW; Abedi-Valugerdi, M. 2010. "28-Day dietary exposure of mice to a low total dose (7mg/kg) of perfluorooctanesulfonate (PFOS) alters neither the cellular compositions of the thymus and spleen nor humoral immune responses: Does the route of administration play a pivotal role in PFOS-induced immunotoxicity?" *Toxicology* 267(1-3):132-139. doi: 10.1016/j.tox.2009.10.035.

Rothman, KJ; Greenland, S; Lash, TL. 2008. *Modern Epidemiology (Third Edition).* Lippincott Williams & Wilkins, Philadelphia, PA, 758p. Accessed on June 27, 2016 at http://site.ebrary.com/lib/gradientcorp/reader.action?docID=10825449.

Ryu, MH; Jha, A; Ojo, OO; Mahood, TH; Basu, S; Detillieux, KA; Nikoobakht, N; Wong, CS; Loewen, M; Becker, AB; Halayko, AJ. 2014. "Chronic exposure to perfluorinated compounds: Impact on airway hyperresponsiveness and inflammation." *Am. J. Physiol. Lung Cell. Mol. Physiol.* 307(10):L765-L774. doi: 10.1152/ajplung.00100.2014.

Shea, BJ; Grimshaw, JM; Wells, GA; Boers, M; Andersson, N; Hamel, C; Porter, AC; Tugwell, P; Moher, D; Bouter, LM. 2007. "Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews." *BMC Med. Res. Methodol.* 7:10. doi: 10.1186/1471-2288-7-10.

Singh, TS; Lee, S; Kim, HH; Choi, JK; Kim, SH. 2012. "Perfluorooctanoic acid induces mast cell-mediated allergic inflammation by the release of histamine and inflammatory mediators." *Toxicol. Lett.* 210(1):64.70. doi: 10.1016/j.toxlet.2012.01.014.

Stephens, ML; Betts, K; Beck, NB; Cogliano, V; Dickersin, K; Fitzpatrick, S; Freeman, J; Gray, G; Hartung, T; McPartland, J; Rooney, AA; Scherer, RW; Verloo, D; Hoffmann, S. 2016. "The emergence of systematic review in toxicology." *Toxicol. Sci.* 152(1):10-16.

Vetvicka, V; Vetvickova, J. 2013. "Reversal of perfluorooctanesulfonate-induced immunotoxicity by a glucan-resveratrol-vitamin C combination." *Orient. Pharm. Exp. Med.* 13(1):77-84. doi: 10.1007/s13596-013-0105-7.

Yang, Q; Abedi-Valurgerdi, M; Xie, Y; Zhao, XY; Moller, G; Nelson, BD; DePierre, JW. 2002. "Potent suppression of the adaptive immune response in mice upon dietary exposure to the potent peroxisome proliferator, perfluorooctanoic acid." *Int. Immunopharmacol.* 2(2-3):389-397.

Zheng, L; Dong, GH; Jin, YH; He, QC. 2009. "Immunotoxic changes associated with a 7-day oral exposure to perfluorooctanesulfonate (PFOS) in adult male C57BL/6 mice." *Arch. Toxicol.* 83(7):679-689.