

National Toxicology Program

**Peer Review of the Draft NTP Monograph on the Systematic
Review of Long-Term Neurological Effects Following Acute
Exposure to the Organophosphorus Nerve Agent Sarin**

February 4, 2019

**National Institute of Environmental Health Sciences
Research Triangle Park, NC**

Peer-Review Report

Table of Contents

1. Attendees	4
2. Introductions and Welcome.....	5
3. Public Comments	5
3.1. Written Public Comments	5
3.2. Oral Public Comments	5
4. Peer Review of the Draft NTP Monograph on Systematic Review of Long-Term Neurological Effects Following Acute Exposure to the Organophosphorus Nerve Agent Sarin.....	5
4.1. Introduction to the Draft NTP Monograph	5
4.1.1. Presentation.....	5
4.1.2. Peer-Review Comments and Panel Discussion on Introduction to the Draft NTP Monograph	6
5. Peer Review of Health Effect Areas.....	7
5.1. Changes in Cholinesterase Levels.....	7
5.1.1. Presentation.....	7
5.1.2. Peer-Review Comments and Panel Discussion on Changes in Cholinesterase Levels, Animal Data	7
5.1.3. Peer-Review Comments and Panel Discussion on Changes in Cholinesterase Levels, Human Data	9
5.2. Visual and Ocular Effects	10
5.2.1. Presentation.....	10
5.2.2. Peer-Review Comments and Panel Discussion on Visual and Ocular Effects, Animal Data.....	11
5.2.3. Peer-Review Comments and Panel Discussion on Visual and Ocular Effects, Human Data	11
5.3. Learning, Memory, and Intelligence Effects	12
5.3.1. Presentation.....	12
5.3.2. Peer-Review Comments and Panel Discussion on Learning, Memory, and Intelligence Effects, Animal Data.....	13
5.3.3. Peer-Review Comments and Panel Discussion on Learning, Memory, and Intelligence Effects, Human Data	16
5.4. Nervous System Morphological and Histological Changes	17
5.4.1. Presentation.....	17
5.4.2. Peer-Review Comments and Panel Discussion on Nervous System Morphological and Histological Changes, Animal Data	18
5.4.3. Peer-Review Comments and Panel Discussion on Nervous System Morphological and Histological Changes, Human Data	19
5.5. Other Outcomes that Did Not Reach Hazard Conclusions	20

5.5.1.	Presentation.....	20
5.5.2.	Peer-Review Comments and Panel Discussion on Other Outcomes that Did Not Reach Hazard Conclusions	21
5.6.	Integration of Animal and Human Evidence for Reaching Hazard Categorization	22
6.	Closing Remarks on the Draft NTP Monograph.....	23
7.	Approval of the Peer Review Report by the Chair of the Peer Review Panel	24

1. Attendees¹

Peer-Review Panel

Chair: Pam Factor-Litvak, Columbia University Medical Center

Frédéric Baud, Université Paris Diderot, Assistance Publique – Hôpitaux de Paris (by WebEx)

John Beard, Brigham Young University (by WebEx)

Peter Blain, Newcastle University (by WebEx)

Michelle Block, Indiana University School of Medicine (by WebEx)

Arik Eisenkraft, The Hebrew University Faculty of Medicine (by WebEx)

Lawrence Engel, University of North Carolina at Chapel Hill (by WebEx)

Virginia Moser, Private Neurotoxicology Consultant (by WebEx)

National Toxicology Program Board of Scientific Counselors Liaison

Kenneth McMartin, Louisiana State University Health Sciences Center (by WebEx)

National Institute of Environmental Health Sciences Staff

Brian Berridge

Andrew Rooney

John Bucher

Mary Wolfe

Elizabeth Maull

Other Federal Agency Staff

David Jett, National Institute of Neurological Disorders and Stroke (by WebEx)

Contract Support Staff

Robyn Blain, ICF (by WebEx)

Ernie Hood, Bridport Services

Canden Byrd, ICF

Chris Sibrizzi, ICF

Pamela Hartman, ICF

Catherine Smith, ICF

¹The meeting was webcast. Individuals who viewed the webcast are not listed except as noted.

2. Introductions and Welcome

The National Toxicology Program (NTP) convened a peer-review panel for the *Draft NTP Monograph on Systematic Review of Long-Term Neurological Effects Following Acute Exposure to the Organophosphorus Nerve Agent Sarin* on February 4, 2019 via webcast. Dr. Pam Factor-Litvak served as chair. Dr. Kenneth McMartin viewed the webcast as the NTP Board of Scientific Counselors liaison. Representing NTP were Drs. Brian Berridge, John Bucher, Elizabeth Maull, Andrew Rooney, and Mary Wolfe. Dr. Maull served as the Designated Federal Official.

Dr. Factor-Litvak called the meeting to order at 9:00 a.m., welcomed everyone to the meeting, and asked all attendees to introduce themselves. Dr. David Jett from the National Institute of Neurological Disorders and Stroke, Director of the National Institutes of Health Countermeasures Against Chemical Threats (CounterACT) program added his welcome and provided background information about CounterACT. Dr. Berridge welcomed all participants to the meeting. Dr. Maull read the conflict of interest policy statement and briefed the attendees on meeting logistics. Dr. Factor-Litvak informed the panel and the audience of the format for the peer review.

3. Public Comments

3.1. Written Public Comments

No written public comments on the draft monograph were received.

3.2. Oral Public Comments

No requests for oral public comments on the draft monograph were received.

4. Peer Review of the Draft NTP Monograph on Systematic Review of Long-Term Neurological Effects Following Acute Exposure to the Organophosphorus Nerve Agent Sarin

4.1. Introduction to the Draft NTP Monograph

4.1.1. Presentation

Dr. Andrew Rooney, Acting Director of the Division of the NTP Office of Health Assessment and Translation (OHAT), presented an introduction to the monograph. Background information on sarin included well known short-term health effects of acute exposure as well as less well characterized long-term neurological effects of human sarin exposure. Sarin was nominated by CounterACT, which requested that NTP conduct a systematic review of the evidence for long-term neurological effects of exposure to sarin to inform the need to develop therapeutics.

Dr. Rooney described the OHAT approach to conducting systematic reviews and the specific process used in the sarin systematic review. The stepwise methods identify, evaluate, and integrate evidence from animal and human studies to reach hazard conclusions on whether sarin is associated with long-term neurological effects. NTP's confidence ratings translated directly

into level of evidence conclusions that also considers the direction of the effect (i.e., confidence that the evidence supports a health effect or no effect). The highest level of evidence conclusions for each time period was used to develop initial hazard identification conclusions. Hazard identification conclusions were developed for three post-exposure time periods: initial (>24 hours to 7 days after exposure), intermediate (8 days to 1 year after exposure), and extended (>1 year after exposure). Four main health effect categories were identified: changes in cholinesterase levels; visual and ocular effects; learning, memory, and intelligence effects; and nervous system morphological and histological changes. Other outcomes were considered, with data included in Appendix 4 of the monograph.

4.1.2. Peer-Review Comments and Panel Discussion on Introduction to the Draft NTP Monograph

Dr. Arik Eisenkraft said that he appreciated NTP's thorough work. He felt that data on health effects from other organophosphates should have been included. Although the focus on cholinesterase was appropriate, he observed that other neurotransmitter pathways were ignored. Dr. Rooney responded that the process of problem formulation was challenging, and in order to meet the needs of CounterACT, it was decided to focus on sarin only.

Dr. John Beard appreciated Dr. Rooney's clarification of the treatment of non-English studies.

Dr. Frédéric Baud questioned whether the time periods in the experimental animal studies translated well to human time periods. Dr. Rooney noted that for the initial time period, NTP used equivalent time periods between the human and non-human animals. However, for the intermediate time period, animal studies with time periods up to 90 days were considered equivalent to a human time period up to 1 year given the shorter life span of rodents compared to humans. He said NTP would appreciate suggestions for any modification of the approach, and would attempt to clarify the text regarding animal time periods in the monograph.

Dr. Peter Blain noted the difficulty of conducting a review on a nerve agent due to the limited number of studies in the public arena. He wondered if there was any way to liaise with appropriate agencies to access classified data relevant to the monograph, or to have agencies comment on the monograph. Dr. Rooney indicated that NTP shared the monograph on an inter-agency level, including the Departments of Defense and Homeland Security and other agencies with access to classified data.

Dr. Rooney presented the peer review panel's charge to:

- 1) Comment on whether the *Draft NTP Monograph on Systematic Review of Long-Term Neurological Effects Following Acute Exposure to the Organophosphorus Nerve Agent Sarin* is technically correct, clearly stated, and objectively presented.
- 2) Vote on whether the scientific evidence from animal studies and from human studies supports the level of evidence conclusions regarding health effects following acute sarin exposure.
- 3) Vote on whether the scientific evidence supports NTP's policy decisions for hazard categorization on long-term neurological effects following acute sarin exposure.

5. Peer Review of Health Effect Areas

5.1. Changes in Cholinesterase Levels

5.1.1. Presentation

Dr. Rooney presented the draft monograph information on changes in cholinesterase levels.

Dr. Rooney described the body of evidence as well as factors that increased or decreased NTP's confidence considerations for both the animal and human studies. The confidence ratings and corresponding level of evidence conclusions in the body of evidence for animal studies were:

- Initial time period (>24 hours to 7 days): moderate confidence, moderate level of evidence
- Intermediate time period (8 days to 1 year): moderate confidence, moderate level of evidence
- Extended time period (>1 year): no confidence rating, inadequate level of evidence

The confidence ratings and corresponding level of evidence conclusions on the body of evidence for human studies were:

- Initial time period (>24 hours to 7 days): high confidence, high level of evidence
- Intermediate time period (8 days to 1 year): very low confidence, inadequate level of evidence
- Extended time period (>1 year): low confidence, inadequate level of evidence

5.1.2. Peer-Review Comments and Panel Discussion on Changes in Cholinesterase Levels, Animal Data

Dr. Peter Blain, first reviewer, noted the small number of studies involved, but said the review of the studies was reasonable and the conclusions were justified. He suggested adding differentiation between butyrylcholinesterase and acetylcholinesterase, as the two enzymes have different kinetics for breakdown of sarin and re-synthesis. He added that the focus on blood did not necessarily reflect the kinetics in the nervous system tissues such as the brain. The monograph could acknowledge that changes in cholinesterase are sometimes viewed as a surrogate marker of exposure and recovery. With those caveats delineated in the monograph, he opined that the conclusions reached were perfectly reasonable. Dr. Peter Blain also noted that decreases in cholinesterase brain levels in the initial time period may be relevant to memory and cognition symptoms in human cases.

Dr. Virginia Moser, second reviewer, indicated that the Damodoran 2003 was not included or described in the discussion on cholinesterase inhibition; if it was not discussed, it should not be listed as one of the studies. She disagreed with downgrading some studies based on sample size or failure to blind assessors to treatments. Dr. Moser indicated that the sample sizes for the rodent studies were actually quite high for cholinesterase studies. For those studies where the assessments were the actual cholinesterase assays themselves, she suggested that blinding of the technician would not be of concern. In the initial time period, while NTP concluded that only a moderate confidence level was appropriate due to differences across studies, Dr. Moser noted

strikingly similar cholinesterase data across the various experiments. Of the nine rodent studies that measured cholinesterase in the initial time period, eight of them recorded decreases at one, two, or three days. The only exception was the Bansal 2009 study, which was the only study to use isopropanol as the vehicle; she was not convinced that isopropanol would not alter sarin absorption. Dr. Moser noted the only two studies that did not report cholinesterase inhibition at seven days was Bansal 2009 and RamaRao 2011, which used female rats. Even with the lack of information on the risk of bias issues of randomization and chemical purity, she felt that downgrading for risk of bias was unwarranted. Dr. Moser suggested NTP should consider a high level of evidence conclusion in the context of the consistent evidence of effects seen during the initial time period. For the intermediate time period, she agreed with moderate confidence in the database. Citing dose response limitations in the two studies discussed at the longer time points, she stated that upregulation should be downplayed as a factor in the intermediate exposure. She noted that there were no animal studies in the extended time period, and thus had no comments on NTP's level of evidence conclusion.

Dr. Rooney appreciated Dr. Peter Blain's thoughts about adding data on the re-synthesis kinetics and linking cholinesterase effects on different brain regions to the other monograph sections such as morphology. Responding to Dr. Moser's comments, Dr. Rooney indicated that he would check on the use of isopropanol as a vehicle and review the upregulation issue. He noted that risk-of-bias assessment is used in systematic reviews to address study quality and transparently report where there might be concerns in study design or conduct that could impact the results. With regard to blinding, he noted that even in more equipment-based procedures, there remains the potential for researchers to intentionally or unintentionally bias the results and therefore there is greater certainty in results that are measured where researchers are blind to treatment level. He noted that the panel is free to disagree with NTP's assessments of risk of bias and provide alternate study quality ratings.

Dr. Eisenkraft agreed with Dr. Moser on her assessment of the body of evidence for the initial time period. He suggested inclusion of descriptions of which studies were performed on rats, and which were performed on non-rodent animals, to define how animal model may impact risk-of-bias ratings.

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the initial time period in animal studies. Dr. Peter Blain moved to accept the level of evidence conclusion as written, Dr. Michelle Block seconded. The panel voted 5 yes, 2 no, 0 abstentions. Drs. Eisenkraft and Moser explained that their no votes ensued from their belief that the conclusion should have been a high level of evidence.

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the intermediate time period in animal studies. Dr. Moser moved to accept the level of evidence conclusion as written, Dr. Peter Blain seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the extended time period in animal studies. Dr. Block moved to accept the level of evidence conclusion as written, Dr. Peter Blain seconded. The panel voted 7 yes, 0 no, 0 abstentions.

5.1.3. Peer-Review Comments and Panel Discussion on Changes in Cholinesterase Levels, Human Data

Dr. Eisenkraft, first reviewer, thought it would be important to add the Tokyo, Japan, events, where sarin of relatively low quality was deployed. He thought that a group of experts needed to be involved in the research, especially given the current use of sarin in the Middle East. Including “time-to-treatment” data should be considered for the monograph. Dr. Eisenkraft agreed with the level-of-evidence conclusions.

Dr. Lawrence Engel, second reviewer, noted that it is nearly impossible to accurately assess the level of exposure. While the best exposure surrogate is often cholinesterase, cholinesterase per se is not a measure of exposure. Recognizing that as a limitation, he agreed with the assessment in the early time period. However, Dr. Engel disagreed with the very low confidence rating for the intermediate time period. The strong effects of butyrylcholinesterase and acetylcholinesterase inhibition in the blood should offset some of the limiting factors.

Dr. Rooney said he appreciated the thoughtful comments from both reviewers. He said he would edit the text to address Dr. Eisenkraft’s comments. He thanked Dr. Engel for his suggestion to upgrade the intermediate time period conclusion from very low to low.

Dr. Moser agreed with Dr. Engel’s suggestion regarding the confidence rating for the intermediate time period.

Dr. Beard commented that “case series” should be identified as “cohort studies” since humans are identified by cholinergic signs and symptoms (exposure versus the outcome). Dr. Beard added that a confounder should be defined as a third variable that is associated with the exposure and the outcome, and it cannot be affected by a prior exposure. Regarding the Tokyo and Matsumoto attacks in Japan, terrorist attacks seem like random events, and Dr. Beard struggled to identify determinants of the exposure and, therefore, likely confounders. He questioned whether downgrading level of evidence because of the risk of bias in terms of confounding was warranted in some cases, as it would be difficult to determine what to adjust for as the confounder.

Dr. Factor-Litvak indicated that confounding is easily confused with effect modification, which means that the reported associations might differ by some characteristic. Thus, what to control for can be very tricky, particularly in considering factors such as sex, which should be considered an effect modification variable. She also commented that treatment for the acute episode might reduce the later outcomes, or bias results toward the null. The issue is how the direction of bias comes in play with treatment.

Dr. Baud noticed six or seven studies referencing cholinesterase inhibition at 13 days following exposure, and therefore thought that an inadequate level of evidence for the intermediate time period was not adapted to the data.

Based on the discussion, Dr. Rooney indicated that he would examine the issue of confounding versus effect modification and re-examine the risk-of-bias ratings to see how treatment effects were addressed. Cholinesterase levels were not the only indicator of exposure; individuals were assumed exposed due to their presence at the attack sites. Confidence ratings were not influenced by how studies were described (i.e. if a study was described as “case-series”) but rather on the presence of study design characteristics such as whether exposure preceded outcome assessment.

Dr. Peter Blain noted that in one study, effects were seen long after cholinesterase levels had returned to normal, so cholinesterase is just an effect marker, and not necessarily a direct marker of an adverse effect. Dr. Factor-Litvak suggested that perhaps cholinesterase is an intermediate marker on a pathway to another outcome.

Dr. Rooney asked the panel for clarification: should cholinesterase be considered as only a biomarker of exposure? If so, does that preclude its use in the conclusions? Or can NTP provide additional language in the text to better describe the situation?

Dr. Engel commented that cholinesterase should be listed and interpreted as an effect. However, what is being measured in the human studies is a surrogate of the effect and should be clearly communicated as such. Dr. Factor-Litvak agreed that intermediate markers are important.

Dr. Factor-Litvak called for a motion on the conclusion of a high level of evidence for the initial time period in human studies. Dr. Engel moved to accept the level of evidence conclusion as written, Dr. Peter Blain seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the intermediate time period in human studies. Dr. Engel moved to change the level of evidence conclusion to low, Dr. Moser seconded. The panel voted 7 yes, 0 no, 0 abstentions to change the level of evidence conclusion to “low.”

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the extended time period in human studies. Dr. Peter Blain moved to accept the level of evidence conclusion as written, Dr. Engel seconded. The panel voted 7 yes, 0 no, 0 abstentions.

5.2. Visual and Ocular Effects

5.2.1. Presentation

Mr. Chris Sibrizzi from ICF presented the draft monograph information on visual and ocular effects.

Mr. Sibrizzi described the body of evidence as well as factors that increased or decreased NTP’s confidence considerations for both the animal and the human studies. The confidence ratings and corresponding level of evidence conclusions on the body of evidence for animal studies were:

- Initial time period (>24 hours to 7 days): moderate confidence, inadequate level of evidence
- Intermediate time period (8 days to 1 year): very low confidence, inadequate level of evidence
- Extended time period (>1 year): very low confidence, inadequate level of evidence

The confidence ratings and corresponding level of evidence conclusions for the body of evidence for human studies were:

- Initial time period (>24 hours to 7 days): moderate confidence, moderate level of evidence
- Intermediate time period (8 days to 1 year): moderate confidence for 2 cross-sectional studies and very low confidence for 8 case reports/series, moderate level of evidence

- Extended time period (>1 year): very low confidence, inadequate level of evidence

Dr. Factor-Litvak asked if it is possible to misclassify a poorly done study with some evidence for an association, or no evidence for an association as “inadequate” rather than “low” level of evidence. Dr. Rooney indicated that the NTP methodology would specifically indicate that the level of evidence was inadequate to determine if there was an effect rather than stating that nothing was reported.

5.2.2. Peer-Review Comments and Panel Discussion on Visual and Ocular Effects, Animal Data

Dr. Block, first reviewer, commented that the analysis was appropriate and agreed with the final level of evidence conclusions. However, she commented that the entry of “unexplained inconsistency” (Table 10) for an initial study was more a reflection of insufficient data. Dr. Block did not consider the confidence conclusion for the initial time period to be moderate, but agreed that the level of evidence conclusion was inadequate because of the insufficient data. In addition, she thought that one of the studies included in the extended time period should technically be classified as belonging to the intermediate period.

Dr. Moser, second reviewer, concurred with Dr. Block’s analysis and had nothing to add.

Dr. Rooney responded that NTP would consider a change in the confidence call from moderate to low for the initial time period. Regarding time period classifications, animal studies of >90 days were included in the extended time period. Clarification would be provided in the revised monograph.

Dr. Eisenkraft pointed out that the 2017 Egoz *et al.* study was mischaracterized in the monograph. Dr. Rooney agreed and indicated that the text would be revised to reflect his comment.

Dr. Baud suggested that an inadequate level of evidence for the initial period was too conservative and should be upgraded to low.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the initial time period in animal studies. Dr. Baud moved to change the level of evidence conclusion to low, and no one seconded. Dr. Block moved to accept the level of evidence conclusion as inadequate, Dr. Moser seconded. The panel voted 5 yes, 2 no, 0 abstentions. Drs. Baud and Eisenkraft both considered the inadequate level of evidence too conservative.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the intermediate time period in animal studies. Dr. Block moved to accept the level of evidence conclusion as written, Dr. Moser seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the extended time period in animal studies. Dr. Engel moved to accept the level of evidence conclusion as written, Dr. Peter Blain seconded. The panel voted 7 yes, 0 no, 0 abstentions.

5.2.3. Peer-Review Comments and Panel Discussion on Visual and Ocular Effects, Human Data

Dr. Baud, first reviewer, mentioned a difference in the number of cases in the different time periods as a cause for concern.

Dr. Beard, second reviewer, reiterated his earlier comment about case series versus cohorts, although it would not affect the level of evidence conclusion. Referencing Dr. Factor-Litvak's previous comments on confounders and effect measure modifiers, Dr. Beard recommended upgrading the level of evidence from inadequate to low for the extended time period, as the Nakajime 1999 study had controlled for confounding.

Dr. Rooney asked Dr. Baud to elaborate on his concern about the number of cases in the different time periods. Dr. Baud said he was concerned that the initial time period only had about 300 patients, whereas the intermediate and extended time periods had closer to 3,000. Dr. Rooney said NTP would go back and look at the studies involved.

Referring to the visual and ocular responses, Dr. Peter Blain said there was a common thread running through the monograph suggesting an effect following exposure.

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the initial time period in human studies. Dr. Beard moved to accept the level of evidence conclusion as written, Dr. Peter Blain seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the intermediate time period in human studies. Dr. Engel moved to accept the level of evidence conclusion as written, Dr. Block seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the extended time period in human studies. Dr. Beard moved to change the level of evidence conclusion to "low." Dr. Engel seconded. The panel voted 7 yes, 0 no, 0 abstentions to change the level of evidence conclusion from inadequate to low for the extended time period.

5.3. Learning, Memory, and Intelligence Effects

5.3.1. Presentation

Dr. Robyn Blain from ICF presented the draft monograph information on learning, memory, and intelligence effects.

Dr. Robyn Blain and Mr. Sibrizzi described the body of evidence as well as factors that increased or decreased NTP's confidence considerations for the animal and human studies, respectively. The confidence ratings and corresponding level of evidence conclusions for the body of evidence for animal studies were:

- Initial time period (>24 hours to 7 days): moderate confidence, moderate level of evidence
- Intermediate time period (8 days to 1 year): moderate confidence, moderate level of evidence
- Extended time period (>1 year): low confidence, low level of evidence

The confidence ratings and corresponding level of evidence conclusions for the body of evidence for human studies were:

- Initial time period (>24 hours to 7 days): no confidence rating, inadequate level of evidence

- Intermediate time period (8 days to 1 year): low confidence for 1 cross-sectional study and low confidence for 2 case reports, low level of evidence
- Extended time period (>1 year): moderate confidence for 2 cross-sectional studies and very low confidence for 2 case series, moderate level of evidence

Responding to a question posed by Dr. Factor-Litvak, Dr. Robyn Blain indicated that the results from human studies in the monograph were not stratified by sex. Dr. Factor-Litvak further asked if intelligence was considered or if memory and executive function were measured in any of the human studies. Dr. Robyn Blain cited one case report of an Army sergeant exposed to sarin who underwent IQ testing. He had difficulty remembering numbers, a memory effect.

Dr. Peter Blain was concerned about translating neurobehavioral animal studies to humans as done in the monograph. He pointed out that humans have higher executive function than what is achievable in non-humans. He asked if the most common change between the animal and human studies was memory. Dr. Robyn Blain confirmed that the most common neurobehavioral effect observed in the animal and human studies involved memory.

5.3.2. Peer-Review Comments and Panel Discussion on Learning, Memory, and Intelligence Effects, Animal Data

Dr. Block, first reviewer, agreed with most of the animal data. However, she was somewhat concerned about the need to combine several different types of tests into one overarching category. She thought the confidence rating for the intermediate time period was low, but could be convinced that it should be moderate. She noted conflicting analyses, particularly in the rat discussion.

Dr. Moser, second reviewer, was concerned about the assessments and interpretations of the studies. She felt that in terms of the methods description and evaluation, the data presentation and study quality factors need to be looked at again. Dr. Moser agreed with a downgrade in confidence rating. However, she thought that some of the information was incorrect. While NTP stated that only the Grauer 2008 paper reported how the animals were randomized to treatment groups, Dr. Moser was unable to find that information in the paper. More importantly, three other studies (Genovese 2009, Pearce 1999, and Muggleton 2003) clearly assigned the animals to the treatment groups to balance performance factors – which she suggested is a standard and appropriate way to assign pre-trained animal to different groups to protect against pre-existing differences in their performance. Regarding outcome assessments, Dr. Moser indicated that eight of nine studies used automated equipment, while the monograph stated that only five studies used automated equipment. An automated visual tracking system that collects and analyzes all data was used for both water maze studies (Allon 2011 and Grauer 2008). Noting a comment in the heat map stating that it was unclear if a particular system analyzed the data, Dr. Moser confirmed that it did, citing her 20 years of experience. She indicated that the Wolthuis 1995 study also used a completely automated computer system. However, there was no information on how the data were collected for the Kassa 2001b T-maze method. It can be assumed that the data were collected by the observer, but it is unknown if the observer was blinded.

The summary for the overall discussion states that all the learning and memory tests used acceptable methods, but Dr. Moser suggested this is not the case. According to Dr. Moser, the water maze procedure used by Allon 2011 and Grauer 2008 studies failed to conform to the

standard method originally described by Morris for the Morris Water Maze. Additionally, the Grauer 2008 study used a non-standard approach for putting the rats on the platform and changed the position of the platform daily to test reference memory. While Grauer 2008 stated that their procedure did not impact the outcome, Dr. Moser stated that many other studies have shown otherwise. The cues can alter the cognitive processes that are used by the animals to complete the task. So this cannot be considered a true assessment of reference memory. The only dependent variable in the Kassa study (Kassa 2001b) with the T-maze was time to reach the goal box, and this was greatly impacted by motor changes. Without any other data on motor functioning, the increased latency cannot be considered a clear cognitive effect. Furthermore, T-mazes are typically used for positional discrimination studies, either using a spontaneous or delayed alternation method, which was not the case in the Kassa study (Kassa 2001b). Dr. Moser indicated that in fact, the Kassa 2001b study used arms that were different colors, so instead of evaluating positional or spatial discrimination, it is actually evaluating cued discrimination. She suggested that these points are all very important for understanding the data obtained from these studies.

In the initial time period, Dr. Moser continued, NTP concluded that the seven studies represented consistent effects. She concluded that these studies do not actually represent evidence of cognitive effects in rats and marmosets, but instead provide inadequate evidence of effects, or possibly even evidence of no effect. The Kassa studies (Kassa 2001b; Kassa 2002; Kassa 2004) used the T-maze and the Y-maze. One study reported an effect on latency, which is not a clear cognitive effect. The 2002 and 2004 Kassa studies used a Y-maze, and reported increased latency to the goal box. The studies, which also measured arm entries, reported no increase in the entry error, that is, entering into the wrong arm. She indicated that the error rate is a better measure of memory compared to latency, which is affected by motor function. For this reason, Dr. Moser said she interpreted the data on latency as a measure of motor function, not a specific cognitive effect. She noted that the 2002 and 2004 Kassa studies both report data that are almost exactly the same, although the figures are graphed differently and with different scales.

Dr. Moser considered the Genovese 2009 study misrepresented in the monograph. The statistical analysis in the paper does not support the monograph statement that total errors were significantly altered in the first block, and that there were no overall significant dose effects and no dose-by-block interactions. She indicated that no further analyses should be conducted in the absence of an overall significant effect; however, the Genovese 2009 author proceeded with the step-down analysis and found effects in the first block. NTP's assumption that this was meaningful directly contradicts the authors' summary, who stated in their discussion that no differences existed on measures of accuracy (working and reference errors) and no difference existed in completion time between rats in the first single exposure.

Dr. Moser continued, and indicated that although the monograph summary describes the results of the monkey studies as inconsistent, the studies very consistently found no effect. The Wolthuis 1995 and Muggleton 2003 studies showed acute effects of treatment that did not persist past the day of exposure, and the Pearce 1999 study reported no negative effects on behavior. On the other hand, both the Muggleton 2003 and Pearce 1999 studies reported improved error rates and improvement on certain components of discrimination sequences in the days after dosing, so there was consistent evidence of no impairment to learning or memory, and some suggestion of improved performance in the initial time period.

These studies consistently demonstrate a lack of effects, except for endpoints that are based on motor functions rather than specifically learning and memory, as well as potentially improved function on some measures. None of the data showed a true dose response. Given these questionable findings, Dr. Moser rated these studies with a low final confidence rating.

The intermediate time period effects are based mostly on some of the same papers that provided the data for the initial time period, so Dr. Moser indicated she would not go over those data again. Two additional papers (Allon 2011 and Grauer 2008) relevant for this time period came from the same laboratory. Both studies used the same non-standard procedure for the water maze with only the Grauer 2008 paper reporting little improvement in latency between trials on the same day and across days in treated rats. While this clearly suggests effects on learning and memory, severe toxicity was also observed, including convulsions and high mortality after exposure. Convulsions produced by a variety of insults can cause hippocampal damage, which, in turn, can impact maze performance. Therefore, Dr. Moser stated it is unclear if the water maze effects were due to the sarin or if they were the result of the sarin-induced convulsions. The Allon 2011 paper, using a lower, non-lethal, exposure level which produced limited toxicity (10%) in the exposure group, reported no change in water maze behavior, supporting Dr. Moser's assertion that the observed learning and memory effects in Grauer 2008 are more a consequence of the debilitating toxicity, including convulsions. Taken together, Dr. Moser thought there was little evidence of impaired cognitive function in the intermediate time period. The inconsistency of these findings, with the effects on motor function and improved performance, warrant another downgrade. Also, upgrading based on the dose-response is not appropriate in studies where it looks like the data were based on motor function, not cognition.

Dr. Moser's comments on the studies in the extended time period were the same as her comments on the studies in the intermediate time period, as the water maze studies and a single the monkey study were the only ones that evaluated the longer time period. She saw these papers as supporting no effect or being inadequate to support any kind of effect on learning and memory in animals.

Dr. Rooney said he appreciated Dr. Moser's comments and expertise. OHAT's risk of bias procedure calls for contacting study authors for additional details not included in a publication. Based on her comments, NTP would reconsider the risk-of-bias evaluations. He noted that they have tried to separate out motor activity effects from behavioral effects in previous evaluations and would consider edits to the text to address this issue.

Dr. Eisenkraft added that it is sometimes difficult to discriminate between the toxic effects of exposure and hypoxic or convulsion-related effects in the longer time periods. More clinical context should be provided when learning and memory deficits are considered.

Dr. Peter Blain agreed with Dr. Eisenkraft's comments. Higher-level effects are likely to be the result of damage to the hippocampus and limbic structures due to excitotoxic or hypoxic effects, and not necessarily a direct toxic effect of exposure to sarin.

Dr. Factor-Litvak brought up the potential impact of survivor effects as a potential source of bias in the animal studies. Dr. Moser said that learning and memory studies generally avoid consideration of survivor effects; they are not typically included in statistical analysis of the results.

Dr. Block suggested that the response mechanisms would be different for lower dose versus a higher dose in an acute exposure scenario. Longer-term effects could lead to persistent, chronic neuropathological effects.

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the initial time period in animal studies. Dr. Moser moved to change the level of evidence conclusion to “low.” Dr. Block seconded. The panel voted 7 yes, 0 no, 0 abstentions to change the level of evidence conclusion to “low.”

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the intermediate time period in animal studies. Dr. Moser moved to change the level of evidence conclusion to “low.” Dr. Block seconded. The panel voted 7 yes, 0 no, 0 abstentions to change the level of evidence conclusion to “low.”

Dr. Factor-Litvak called for a motion on the conclusion of a low level of evidence for the extended time period in animal studies. Dr. Block moved to accept the level of evidence conclusion as written. Dr. Moser seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Rooney asked the panel to comment on whether the learning and memory testing seen in the studies would be useful in testing therapeutics, given the changes across the board to a “low” level of evidence. Dr. Eisenkraft said that because of the low level of evidence, there must be more studies. Dr. Moser agreed that more studies are needed, with improved methods. Dr. Peter Blain said that higher-animal studies, such as in monkeys and non-human primates, would be more valuable in neurobehavioral assessments than studies in rodents. Dr. Factor-Litvak suggested that NTP recommend a standard protocol for testing, including rodent studies. Dr. Eisenkraft pointed out that some of the studies are deemed to be equivalent to clinical studies that cannot be performed in humans by the U.S. Food and Drug Administration (FDA); the FDA should be involved in helping with study design in such cases.

5.3.3. Peer-Review Comments and Panel Discussion on Learning, Memory, and Intelligence Effects, Human Data

Dr. Beard, first reviewer, repeated his earlier comments about case series versus cohort, and confounding. While he agreed with the NTP’s level of evidence conclusion, he disagreed with the learning and memory outcomes summaries in Table 1 for the studies conducted by Miyaki in 2005 and Nishiwaki in 2001 and took issue with the treatment of small sample size. Dr. Beard suggested wording changes for both summaries for NTP’s consideration. He also disagreed with the inadequate level of evidence characterization for the intermediate period, as the statement contrasts with other sentences stating that human data provide low evidence.

Dr. Eisenkraft, second reviewer, reiterated the issues regarding long-term psychological effects, which cannot be differentiated between direct sarin exposure and stress from the sarin exposure event.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the initial time period in human studies. Dr. Engel moved to accept the level of evidence conclusion as written. Dr. Beard seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak called for a motion on the conclusion of a low level of evidence for the intermediate time period in human studies. Dr. Beard moved to accept the level of evidence conclusion as written. Dr. Engel seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the extended time period in human studies. Dr. Beard moved to accept the level of evidence conclusion as written. Dr. Eisenkraft seconded. The panel voted 7 yes, 0 no, 0 abstentions.

5.4. Nervous System Morphological and Histological Changes

5.4.1. Presentation

Dr. Robyn Blain from ICF presented the draft monograph information on nervous system morphological and histological effects.

Dr. Robyn Blain and Mr. Sibrizzi described the body of evidence as well as factors that increased or decreased NTP's confidence considerations for animal and human studies, respectively. The confidence ratings and corresponding level of evidence conclusions for the body of evidence for animal studies were:

- Initial time period (>24 hours to 7 days): moderate confidence, moderate level of evidence
- Intermediate time period (8 days to 1 year): moderate confidence, moderate level of evidence
- Extended time period (>1 year): no confidence rating, inadequate level of evidence

The confidence ratings and corresponding level of evidence conclusions for the body of evidence for human studies were:

- Initial time period (>24 hours to 7 days): no confidence rating, inadequate level of evidence
- Intermediate time period (8 days to 1 year): low confidence, inadequate level of evidence
- Extended time period (>1 year): moderate confidence for 1 cross-sectional study and low confidence for 1 cast report, moderate level of evidence

Dr. Factor-Litvak asked if the morphological changes seen in the human data were definitively attributed to sarin exposure. Dr. Robyn Blain explained that while an MRI in the single case report showed changes in the brain, definitive attributions to sarin exposure could not be done with any certainty. That case report would have risk-of-bias concerns. She added that the moderate confidence level was based more on the cross-sectional study.

Dr. Eisenkraft requested that clinical assessments, such as evidence of seizures, be added to the monograph. As Dr. Robyn Blain did not recall reports of seizures, she indicated that she would review the studies.

5.4.2. Peer-Review Comments and Panel Discussion on Nervous System Morphological and Histological Changes, Animal Data

Dr. Peter Blain, first reviewer, noted that few animal studies met the criteria for inclusion. However, there were consistent findings within those studies, particularly regarding limbic structures in the brain and those findings provided evidence for long-term effects. While there were reports of nerve fiber degeneration, fitting with the electrophysiology of the peripheral nervous system, it cannot be determined if this is a direct effect of sarin exposure or due to either a respiratory effect (hypoxia), or to cytotoxic effect resulting from seizure activity.

Dr. Block, second reviewer, opined that even if the effects resulted from an indirect mechanism leading to neurotoxicity, it is still an effect of sarin exposure. She was less enthusiastic about the animal evidence and would downgrade both the initial and intermediate level of evidence conclusions to “low.” There needs to be some method of quantification, even if it is not the same as current methods. Some of the studies do not hit that bar, she observed, and she cited several examples of studies that did not quantify appropriately. Nonetheless, Dr. Block believed that there was evidence to suggest that something is occurring, but only at a gross level. She noted that neuropathology would get worse over time, and there would be ongoing pathology. Dr. Block recommended downgrading the level of evidence rating based on low confidence in the evidence based on low quality of the studies.

In response to a question posed by Dr. Rooney, Dr. Block replied that there was not any specific aspect of the OHAT method that caused her to downgrade the level of evidence; rather the methods used in the studies had been modified over time because they had much bias. Dr. Block indicated that it is important to note the level of assay stringency reflects the time when the study was conducted.

Dr. Peter Blain, speaking from clinical experience, cited organophosphate-related hypoxia and seizure activity causing long-term brain damage. Quantification does not necessarily conform with clinical experience as the effects are not necessarily the result of a direct toxic mechanism of sarin itself; it is the secondary effects, such as damage to the brain, that are of the most concern.

In response to a question posed by Dr. Factor-Litvak, Dr. Robyn Blain indicated that none of the investigators had stratified by levels of hypoxia or convulsions in the animals.

Dr. Eisenkraft stated that most of the models did not represent a realistic scenario, as neither oxygen nor ventilation was provided to the animals, which would be highly unacceptable in humans. Also, the exposure levels in many of the studies were extremely high. Dr. Eisenkraft was unsure if the long-term effects observed were direct or indirect effects of sarin, although it does not matter much, because the effects must be treated. He hoped that one of the things to emerge from the review would be protocol guidelines for future animal studies.

Dr. Peter Blain suggested that a higher animal model would be more advantageous to use, for a more realistic outcome.

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the initial time period in animal studies. Dr. Eisenkraft moved to accept the level of evidence conclusion as written. Dr. Peter Blain seconded. The panel voted 5 yes, 2 no, 0 abstentions. Dr. Block explained her no vote as stemming from her belief that the conclusion should have been

“low”, based on the fact that the studies were not stringent. Dr. Moser also voted no, for similar reasons.

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the intermediate time period in animal studies. Dr. Eisenkraft moved to accept the level of evidence conclusion as written. Dr. Peter Blain seconded. The panel voted 5 yes, 2 no, 0 abstentions. Dr. Block explained her no vote, reiterating the conclusion should have been “low,” because the studies were not stringent. Dr. Moser also voted no, for similar reasons.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the extended time period in animal studies. Dr. Block moved to accept the level of evidence conclusion as written. Dr. Peter Blain seconded. The panel voted 7 yes, 0 no, 0 abstentions.

5.4.3. Peer-Review Comments and Panel Discussion on Nervous System Morphological and Histological Changes, Human Data

Dr. Baud, first reviewer, indicated that there is a very limited body of knowledge, with no studies in the initial time period and only one case report for a single U.S. military member for the intermediate time period. That patient underwent MRI and PET scans, both of which were normal for the brain and the spine. There were two studies in the extended time period: one case report and one cross-sectional study, both stemming from the Tokyo, Japan, subway attack. Dr. Baud summarized the two studies and indicated that it was quite difficult to characterize the conclusion for the extended time period as moderate or low.

Dr. Engel, second reviewer, agreed that there was a very limited amount of evidence, with two case reports of one person each and one cross-sectional study of small-to-moderate size. He found it interesting that in the Loh study, the patient showed no effect on the MRI, but did have a substantial decrement in cholinesterase levels. He said he would score that (the intermediate time period) as a very low level of evidence. As to the extended time period, he was curious whether some of the reported effects could be post-traumatic stress disorder (PTSD)-related. However, he agreed with an earlier comment that it should not matter whether the effect is a direct sarin insult or an indirect effect mediated through some other biological mechanism. A different study design may have been able to address that issue. He agreed with the conclusion of moderate level of evidence for the extended time period.

Dr. Rooney indicated that Dr. Baud’s conclusions were clear. He asked Dr. Baud to submit his specific comments on the two case reports and one cross-sectional study in writing so they could be considered in revisions, because his audio was poor during his comments.

Dr. Robyn Blain said that the PTSD issue would arise later in some of the other studies to be considered. It is difficult to tease out what caused the PTSD. Few sarin studies specifically considered PTSD.

Dr. Eisenkraft indicated that there are many similarities between the long-term neurobehavioral and morphological effects of PTSD and either sarin or other organophosphates. He disagreed with a sentence on page 60 of the monograph saying that “given the subject’s symptoms after the exposure, it is likely that the effects are related to the sarin exposure.” He felt that the symptoms were more of an anoxic event. In his opinion, the Himuro 1998 study looked like hypoxic anoxia not directly related to sarin. He said that it can be seen that sarin exposure causes long-term

damage, with a moderate to high level of confidence. It should be decided if the discussion should address direct or indirect effects.

Dr. Peter Blain said that in recent clinical experience with sarin and other organophosphate nerve agents, for patients that survived, their physiology supported evidence that hypoxia and seizure activity were prevented. Thus, these patients can survive with minimal residual damage by controlling hypoxia and seizure activity.

Dr. Factor-Litvak asked if the evidence suggests a dose-response relationship. Dr. Peter Blain replied that people with low-level organophosphate exposures can and do survive. Dr. Eisenkraft indicated that if the seizures are not stopped as early as possible, they will become more severe and more treatment-resistant, citing evidence from both sarin and other organophosphate nerve agents. He said he was not aware of an accurate dose-response curve with nerve agents for humans.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the initial time period in human studies. Dr. Engel moved to accept the level of evidence conclusion as written. Dr. Eisenkraft seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak called for a motion on the conclusion of an inadequate level of evidence for the intermediate time period in human studies. Dr. Baud moved to accept the level of evidence conclusion as written. Dr. Engel seconded. The panel voted 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak called for a motion on the conclusion of a moderate level of evidence for the extended time period in human studies. Dr. Eisenkraft moved to accept the level of evidence conclusion as written. Dr. Baud seconded. The panel voted 7 yes, 0 no, 0 abstentions.

5.5. Other Outcomes that Did Not Reach Hazard Conclusions

5.5.1. Presentation

Dr. Rooney presented material to the panel regarding other outcomes that did not reach hazard conclusions.

There was inadequate evidence to determine whether there is an association with acute sarin exposure for the following outcomes:

- Sleep disruption
- Anxiety and fear
- Avoidance and depression
- Activity and strength
- Other neurological symptoms
- Electroencephalogram (EEG) data
- Other sensory effects

The studies that failed to reach a level of evidence rating that would support a hazard conclusion were placed in Appendix 4 of the monograph. The confidence ratings were low, very low, or single studies reporting no evidence of health effects, and for some time periods there were no

data. Some of the studies had very serious risk-of-bias concerns. In general, there were very few overlapping endpoints considered across studies.

5.5.2. Peer-Review Comments and Panel Discussion on Other Outcomes that Did Not Reach Hazard Conclusions

Dr. Factor-Litvak opened the discussion for comments from panel members.

Dr. Eisenkraft noted that the first four bullets on the slide (sleep disruption, anxiety and fear, avoidance and depression, and activity and strength) could be combined into a PTSD endpoint. Dr. Rooney said that PTSD had initially been considered as a potential outcome. Dr. Robyn Blain indicated that it was not included as few studies had specifically considered PTSD, and it was not possible to determine if symptoms were caused by PTSD from the stress of the attack or were directly sarin-related.

Dr. Beard said it seemed that in most sections of the report, standard observational studies were separated from case series studies. He cited the Yokoyama 1998c study as an example, and indicated that a low level of evidence was as high as he would go given risk-of-bias concerns.

Dr. Baud suggested that if the anoxic brain damage is considered to be somewhat related, even indirectly, to sarin, then perhaps it should also be considered that sarin may indirectly cause PTSD.

Dr. Factor-Litvak rephrased the question: are the PTSD symptoms more directly related to the event than to sarin? Could PTSD symptoms be considered to be adverse associations due to sarin, including the event and the actual exposure to the compound?

Dr. Baud suggested that the anoxic brain damage be accepted as a downstream effect related to sarin exposure. Dr. Rooney said he had difficulty with the idea that the data could separate whether the sarin exposure resulted in brain damage directly or if sarin would lead to anoxia which would result in brain damage. Similarly, he noted that the attack had two different aspects, one being sarin-related, the other being the attack itself, which is not specific to the chemical exposure; this was the portion that was difficult to separate out to account for PTSD. Dr. Rooney said he would welcome suggestions from the panel on how to communicate that aspect in the monograph.

Dr. Engel commented that distinguishing the effects of events-induced PTSD versus sarin-induced PTSD was an important issue and suggested several potential approaches to studying the problem. He noted that it is an issue of concern that warrants further investigation.

Dr. Factor-Litvak suggested looking at the evidence from the studies that captured sarin plus the attack, and then to compare them to other studies that had attack but without sarin, to see if there were differences in some of the outcomes. Dr. Rooney appreciated the suggestion.

Dr. Eisenkraft raised the issue of cultural differences between nations and peoples.

Dr. Moser commented that the animal studies were even more minimal than described in the monograph, and there is not much data on the endpoints.

Dr. Peter Blain pointed out that sleep disruption has been a feature of anecdotal reports of acute exposures to organophosphates.

In response to a question posed by Dr. Baud, Dr. Rooney replied that there was no particular ranking for the signs and symptoms listed on the slide.

Dr. Factor-Litvak summarized the discussion, indicating general agreement among the panel that the outcomes at present do not reach the level that would warrant a hazard conclusion. With no hazard conclusions, there would be no vote for the section. She asked for any dissent from panel members. Dr. Eisenkraft said he did not disagree, but indicated that it would be important to state the identified issues in order to study them in future investigation. Dr. Rooney said that the identified signs and symptoms would be acknowledged in the monograph.

Dr. Peter Blain requested that data from single fiber electromyography, used to biomonitor patient recovery, be included in the monograph as well.

5.6. Integration of Animal and Human Evidence for Reaching Hazard Categorization

Dr. Rooney presented material to the panel on the process of evidence integration to reach overall hazard conclusions.

He noted that the process includes two stages: an initial hazard conclusion, the result of considering human and animal evidence together, and a final hazard conclusion, which considers the impact of other data such as relevant mechanistic data and biological plausibility of effect. The final hazard conclusions consider whether there is strong support to increase the hazard identification conclusion, or strong opposition to decrease the hazard identification, or no impact on the hazard identification conclusion.

The final hazard conclusions for acute sarin exposure were as follows:

- Initial time period: **Known to be a neurological hazard to humans** based on suppression of cholinesterase.
- Intermediate time period: **Suspected to be a neurological hazard to humans** based on multiple health effects.
- Extended time period: **Suspected to be a neurological hazard to humans** based on multiple health effects.

Dr. Factor-Litvak opened the floor for comments from panel members.

Dr. Eisenkraft indicated that he could not accept the “suspected” hazard conclusion for the intermediate and extended time periods, and recommended that the conclusion be changed to “presumed.” Asked by Dr. Factor-Litvak to elaborate, Dr. Eisenkraft indicated that his suggestion was based on his working experience with other organophosphates, as well as the devastating effects resulting from sarin exposure. Dr. Rooney asked Eisenkraft if he could identify specific evidence that would support a conclusion of “presumed.”

Dr. Factor-Litvak noted that because the evaluation is meant to be transparent to both the scientific and non-scientific communities, and reproducible, the hazard conclusions must be totally based on the published evidence presented during the review.

Dr. Peter Blain suggested including a reference to the relationship between dose and long-term outcomes in the monograph. Dr. Rooney considered that an excellent communication point, such a message would be included.

Since some of the individual health effects for the extended time period were non-classifiable, Dr. Moser asked how a hazard conclusion of “suspected” was determined. Dr. Rooney indicated that any of the time periods could be supported by the body of evidence for a single health effect.

Dr. Baud said that he did not feel that there was sufficient data to support raising the hazard conclusion to a higher level.

Dr. Factor-Litvak said there would be three votes, one for each time period. She called for a vote on the hazard conclusion (“known”) for the initial time period. Dr. Baud moved to accept the conclusion as written, Dr. Engel seconded. The vote was 7 yes, 0 no, 0 abstentions.

Dr. Factor-Litvak asked for a motion to accept the hazard conclusion (“suspected”) for the intermediate time period. Dr. Baud so moved, Dr. Moser seconded. The vote was 6 yes, 1 no, 0 abstentions. Dr. Eisenkraft explained his no vote based on his prior remarks (prior experience with organophosphates, severity of sarin-induced effects).

Dr. Factor-Litvak asked for a motion to accept the hazard conclusion (“suspected”) for the extended time period. Dr. Engel so moved, Dr. Baud seconded. The vote was 6 yes, 1 no, 0 abstentions. Dr. Eisenkraft said he voted no based on the same concerns he had previously expressed.

6. Closing Remarks on the Draft NTP Monograph

Dr. Factor-Litvak asked the panel to comment on the overall organization of the monograph, particularly in terms of clarity and coherent presentation of the information and its synthesis.

Dr. Engel indicated that within the limit of the published data, the conclusion in the monograph is the strongest statement that can be made. Following up on Dr. Eisenkraft’s comments, there would be a case for a stronger case on sarin in the intermediate and extended time periods, based on anecdotal evidence. The biggest issue in reaching a stronger conclusion is the lack of data.

Dr. Factor-Litvak said that it needs to be clear that there is a substantial amount of classified evidence that could not be incorporated in the monograph.

Dr. Eisenkraft thought that the inclusion of raw data from studies would enable better analysis.

Dr. Baud agreed with Dr. Engel that the body of knowledge did not allow the conclusions to go further. There is value in identifying gaps in current knowledge; it would be helpful to highlight them in the monograph.

Dr. Beard asked if NTP could recommend that future studies should include an actual screener for PTSD, so that the issue could be better teased out. Dr. Factor-Litvak felt that that could probably be included as a future recommendation.

Closing the meeting, Dr. Factor-Litvak thanked the reviewers for their hard work and excellent comments. She also thanked the NTP staff and ICF staff for preparing an excellent monograph and a productive meeting. Dr. Rooney thanked Dr. Factor-Litvak for her efforts in chairing the meeting.

Dr. Maull added her thanks to everyone.

Dr. Factor-Litvak adjourned the meeting at 3:00 p.m. EST on February 4, 2019.

7. Approval of the Peer Review Report by the Chair of the Peer Review Panel

This peer review report has been read and approved by the chair of the February 4, 2019 Peer Review of the *Draft NTP Monograph on the Systematic Review of Long-Term Neurological Effects Following Acute Exposure to the Organophosphate Nerve Agent Sarin*.



Pam Factor-Litvak, Ph.D.

Peer Review Panel Chair

Date: May 15, 2019