PEER REVIEW SUMMARY REPORT

External Peer Review Meeting on the *Toxicological Review of 2-Hexanone (CAS No. 591-78-6)*

Prepared for:
Amanda Persad, Ph.D., D.A.B.T.
U.S. Environmental Protection Agency
National Center for Environmental Assessment
109 T.W. Alexander Dr. (B243-01)
Research Triangle Park, NC, 27711

Prepared by:
Versar, Inc.
Contract No. EP-C-07-025
Task Order 26

Peer Reviewers:
Bahie Abou-Donia, Ph.D.
A. John Bailer, Ph.D.
Frederick J. Miller, Ph.D. Fellow ATS
Mohammad I. Sabri, Ph.D.
Alan H. Stern, D.P.H., D.A.B.T.
Bernard Weiss, Ph.D.

June 5, 2008

I. INTRODUCTION

IRIS is an EPA database containing Agency consensus scientific positions on potential adverse human health effects that may result from chronic (or lifetime) exposure, or in select cases less-than-lifetime exposures, to chemicals in the environment. IRIS currently provides health effects information on over 500 chemical substances.

IRIS contains chemical-specific summaries of qualitative and quantitative health information in support of two steps of the risk assessment process, i.e., hazard identification and dose-response evaluation. IRIS information includes a reference dose (RfD) for non-cancer health effects resulting from oral exposure, a reference concentration (RfC) for non-cancer health effects resulting from inhalation exposure, and an assessment of carcinogenicity for both oral and inhalation exposures. Combined with specific situational exposure assessment information, the health hazard information in IRIS may be used as a source in evaluating potential public health risks from environmental contaminants.

The IRIS program, within EPA's National Center for Environmental Assessment (NCEA), developed a Toxicological Review of 2-Hexanone, an assessment of which has not previously appeared on the IRIS database. 2-Hexanone was nominated for IRIS assessment by the Office of Solid Waste and Emergency Response. The draft documents slated for the external peer review contain a chronic oral reference dose, a chronic inhalation reference concentration, and a qualitative cancer assessment.

Peer Reviewers:

Bahie Abou-Donia, Ph.D.
Duke University Medical Center
Durham, NC 27710

A. John Bailer, Ph.D.
Miami University
Oxford, OH 45056

Frederick J. Miller, Ph.D., Fellow ATS
Fred J. Miller & Associates LLC
Cary, NC 27511

Mohammad I. Sabri, Ph.D.
Oregon Health Science University
Portland, OR 97239

Alan H. Stern, D.P.H., D.A.B.T. (Chair)
University of Medicine and Dentistry of New Jersey
Piscataway, NJ 08854

Bernard Weiss, Ph.D.
University of Rochester School of Medicine and Dentistry
Rochester, NY 14642

II. CHARGE TO THE REVIEWERS

The U.S. Environmental Protection Agency (EPA) is seeking an external peer review of the scientific basis supporting the human health assessment of 2-hexanone that will appear on the Agency's online database, the Integrated Risk Information System (IRIS). IRIS is prepared and maintained by the EPA's National Center for Environmental Assessment (NCEA) within the Office of Research and Development (ORD). There is currently no assessment on the IRIS database for the health effects associated with 2-hexanone exposure.

The draft health assessment documents include a chronic Reference Dose (RfD) and a chronic Reference Concentration (RfC). Below are a set of charge questions that address scientific issues in the assessment of 2-hexanone. Please provide detailed explanations for responses to the charge questions.

(A) General Charge Questions:

1. Is the Toxicological Review logical, clear and concise? Has EPA accurately, clearly and objectively represented and synthesized the scientific evidence for noncancer and cancer hazard?

2. Please identify any additional studies that should be considered in the assessment of the noncancer and cancer health effects of 2-hexanone.

3. Please discuss research that you think would be likely to increase confidence in the database for future assessments of 2-hexanone.

4. Please comment on the identification and characterization of sources of uncertainty in sections 5 and 6 of the assessment document. Please comment on whether the key sources of uncertainty have been adequately discussed. Have the choices and assumptions made in the discussion of uncertainty been transparently and objectively described? Has the impact of the uncertainty on the assessment been transparently and objectively described?

Chemical-Specific Charge Questions:

(B) Oral reference dose (RfD) for 2-hexanone

1. A chronic RfD for 2-hexanone has been derived from a 13-month drinking water study (O'Donoghue et al., 1978) in male rats. Please comment on whether the selection of this study as the principal study has been scientifically justified. Has this study been transparently and objectively described in the document? Please identify and provide the rationale for any other studies that should be selected as the principal study.

2. Myofibrillar atrophy of the quadriceps muscle was selected as the critical effect. Please comment on whether the rationale for the selection of myofibrillar atrophy as the critical effect has been scientifically justified. Has this selection been transparently and objectively described in the document? Please provide detailed explanation. Please comment on the selection of myofibrillar atrophy of the quadriceps muscle as the critical effect rather than other endpoints identified in O'Donoghue et al. (1978). Please comment on the selection of myofibrillar atrophy of the quadriceps muscle as compared to the peripheral nerve axonal swelling. Please identify and provide the rationale for any other endpoints that should be considered in the selection of the critical effect.

3. Please comment on the selection of the uncertainty factors applied to the point of departure (POD) for the derivation of the RfDs. For instance, are they scientifically justified? Are they transparently and objectively described in the document?

4. Please comment specifically on the database uncertainty factor of 3 applied in the RfD derivation. Please comment on body of information regarding reproductive, developmental toxicity, and immunotoxicity on 2-hexanone as well as the relevance of toxicity data on n-hexane in the determination of the database uncertainty factor. Please comment on whether the selection of the database uncertainty factor for the RfD has been scientifically justified. Has this selection been transparently and objectively described in the document?

5. Please provide any other comments on the derivation of the RfD.

(C) Inhalation reference concentration (RfC) for 2-hexanone

1. A chronic RfC for 2-hexanone has been derived from a 10-month inhalation study (Johnson et al., 1977) in rats and monkeys. Please comment on whether the selection of this study as the principal study has been scientifically justified. Has this study been transparently and objectively described in the document? Please comment on the use of a 10-month monkey study (Johnson et al., 1977) as opposed to a 72-week rat study (Krasavage and O'Donoghue, 1977). Please identify and provide the rationale for any other studies that should be selected as the principal study.

2. Motor conduction velocity of the sciatic-tibial nerve in monkeys was selected as the critical toxicological effect. Please comment on whether the selection of this critical effect has been scientifically justified. Has this selection been transparently and objectively described in the document? Please provide detailed explanation. Please comment on the use of motor conduction velocity of the sciatic-tibial nerve instead of motor conduction velocity of the ulnar nerve. Please comment on the use of monkey data instead of rat data. Please identify and provide the rationale for any other endpoints that should be considered in the selection of the critical effect.

3. Estimates of the standard deviation of the responses in each dose group are needed to calculate benchmark doses (BMDs) and their corresponding lower confidence

limits (BMDLs). This information was not provided in Johnson et al. (1977), the principal study. Therefore, an indirect method for estimating this missing information on response variability was devised. Please comment on the procedure used to determine the standard deviation. Please comment on the use of digitization as a method to abstract data from Johnson et al. (1977) for the derivation of the inhalation reference concentration.

4. Please comment on the selection of the uncertainty factors applied to the POD for the derivation of the RfCs. Are they scientifically justified? Are they transparently and objectively described in the document?

5. Please comment specifically on the database uncertainty factor of 3 applied in the RfC derivation. Please comment on body of information regarding reproductive, developmental toxicity (including developmental neurotoxicity), and immunotoxicity on 2-hexanone, as well as the comparability and relevance of toxicity data on n-hexane and 2,5-hexanedione in the determination of the database uncertainty factor. Please comment on whether the selection of the database uncertainty factor for the RfC has been scientifically justified. Has the selection of the database uncertainty factor been transparently and objectively described in the document?

6. Please provide any other comments on the derivation of the RfC.

(D) Carcinogenicity of 2-hexanone

1. Under the EPA's 2005 *Guidelines for carcinogen risk assessment* (www.epa.gov/iris/backgr-d.htm), the Agency concluded that there is *inadequate information to assess the carcinogenic potential* of 2-hexanone. Please comment on the scientific justification for the cancer weight of the evidence characterization.

III. GENERAL IMPRESSIONS

Bahie Abou-Donia

The Toxicological Review of 2-Hexanone is an extensive review of literature on the effects of 2-hexanone (methyl butyl ketone) in order to carry out a dose-response assessment to derive human equivalent oral dose and inhalation concentration. A 13-month drinking water study (O'Donoghue et al., 1978) was chosen to drive the 2-hexanone RfD assessment. The critical endpoints evaluated from this study were incidences of myofibrillar atrophy of the quadriceps muscle and the calf muscle in male rats. Uncertainty factors of 300 were used that resulted in an RfD of 2×10^{-1} mg/kg-day. A 10-month inhalation study (Johnson et al., 1977) in monkeys and rats was used to drive 2-hexanone RfC assessment. The sciatic-tibial nerve motor conduction velocity in monkeys was used to drive the RfC. Uncertainty factors of 1,000 were used that resulted in an RfC of 2×10^{-1} mg/m^3.

The report included justification for selecting both studies. It also indicated that the overall confidence in the 2-hexanone RfD and RfC is "medium." This assessment seems to be accurate considering the available database.

Generally, the report clearly presents the data and the conclusions seem sound. Specific comments are listed below.

A. John Bailer

The report on 2-hexanone is well structured. There was no need to explore carcinogenic potency given that there was little information on the carcinogenicity of this chemical. Dose-response data do exist to support the RfD and RfC although assumptions and back-calculations were required to do a BMC calculation to support the RfC construction. Neither data set was overly compelling since these studies involved relatively few dose groups with small numbers of animals exposed in each group (4 dose groups and 10 animals/group for the RfD calculation and 3 dose groups and 8 animals /group for the RfC calculation). The confidence expressed in the report was that the RfD (p. 79) and RfC (p. 80) assessments is "medium" and I believe this is reasonable.

Frederick J. Miller

The information provided in the document does a reasonable job of providing a summary of the studies one might consider in establishing an oral reference dose and inhalation reference concentration. While there are limited data available on the absorption of 2-hexanone, the data clearly show that the parent compound is eliminated from the body over a reasonably short period of time. The likely toxic moiety is the metabolite 2,5-hexanedione, which is itself cleared fairly quickly.

The principal study selected for calculation of an RfD is a 13 month drinking water study in male rats by O'Donoghue et al. (1978). In addition to controls, 3 levels of 2-hexanone

were introduced in the drinking water. The critical endpoints identified in this study were the incidences of myofibrillar atrophy of the quadriceps muscle and the calf muscle. Their selection over other endpoints related to the effects occurring due to axonal atrophy, which has been identified as a good indicator variable for nerve dysfunction. BMD modeling yielded a BMDL of 49.9 mg/kg-day for quadriceps muscle and a slightly higher value of 69.2 mg/kg-day for calf muscle. The presentation of this information was clear and accurate.

For the RfC, an inhalation study (Johnson et al.,1977) using monkeys and rats was chosen as the principal study. Effects on sciatic-tibial nerve MCV was selected as the critical endpoint. A number of points concerning the effect levels and patterns discussed earlier in the document were not brought forward to Section 5.2.1 so the reasoning and defensibility of the selection of the critical endpoint and the use of the monkey data over the rat data are not well defended. RfC values were derived for both rats and monkeys for sciatic-nerve MCV and ulnar MCV. The monkey sciatic-nerve MCV appears to have been picked for three reasons: (1) monkeys are more similar to humans than are rats, and (2) the derived RfC for the sciatic-nerve MCV is slightly lower than that derived for the ulnar-nerve, and (30 the derived RfCs in both species for the sciatic-nerve MCV are lower than those for the ulnar-nerve MCV.

The assumption that the product of exposure concentration and time is constant is problematic for adjusting the BMDL to reflect continuous exposure when the principal study actually used less than continuous exposures. This topic is expanded upon in response to Question C6.

Mohammad I. Sabri

The draft document has reviewed the background information and dose-response assessment of 2-hexanone-toxicity. Both chronic oral reference dose (RfD) and the chronic inhalation reference concentration (RfC) of 2-hexanone have been derived for the Integrated Risk Information System (IRIS). The review is concise, clear and provides a scientific basis for non-cancer and cancer hazard assessment of 2-hexanone. The health effects of 2-hexanone are well presented. The document is clearly written in an understandable manner and cites most, if not all, relevant papers published on the toxic effects of 2-hexanone. The information and conclusions on 2-hexanone appears accurate and sound. The neurological effects are characteristic and sensitive responses of 2-hexanone-toxicity. To-date, there is no evidence that 2-hexanone causes cancer in either in animals or humans. The International Agency for Research on Cancer, The Environmental Protection Agency (EPA) and The Department of Health and Human Services has not classified 2-hexanone as a carcinogen. The overall medium confidence in the principal studies (O'Donoghue, et al., 1978; Johnson et al., 1977) and in the derivation of the RfD and RfC, appears appropriate.

Alan H. Stern

See responses to the following general charge questions.

Bernard Weiss

The document provides a fairly clear description of the available data, along with the reasons for selecting particular studies as the source of the risk assessment evaluation. I was especially concerned, at first, by the choice of an unpublished experiment 30 years old to establish the oral RfD. Given the competition, so to speak, I have to agree with the review panel that, despite its age, this study seemed to offer the most dependable information.

Because some readers may wonder why the agency has undertaken such a review at this time, the final document should observe the presence of 2-hexanone at Superfund sites.

I found the toxicokinetics section clearly presented and a useful introduction to the rest of the document. In this section and elsewhere, the tables were easy to follow.

In a number of instances, detailed below, I thought the document should have commented on certain issues about which some readers might have questions. In some cases, the authors deferred discussion until later sections, but the comments should be placed in the earlier context.

IV. RESPONSE TO CHARGE

(A) General Charge Questions

1. Is the Toxicological Review logical, clear and concise? Has EPA accurately, clearly and objectively represented and synthesized the scientific evidence for noncancer and cancer hazard?

Bahie Abou-Donia

The toxicological review is generally logical and concise. There are some missing studies that are listed in my response to the following charge questions. Also, all available studies indicate that the nervous system is the prime target for 2-heaxnone neurotoxicity, with axonal swelling and degeneration being the hallmark of this neurotoxicity. Incidences of axonal swelling (Table 5) in the brain, spinal cord and peripheral nerves might have been a better indictor of 2-hexanone neurotoxicity than incidences of myofibrillar atrophy.

A. John Bailer

The presentation is structured in a logical order with a build up from chemical information to toxicokinetics (including key metabolites) to hazard identification to dose-response assessments and ultimately, conclusions. The appendices should be improved (see suggestions in the specific comments section).

Frederick J. Miller

The Agency has accurately and objectively laid out the extent of the data available in the literature that can be used to assess the toxicity of 2-hexanone. Deficits in the published studies are identified that may impact the value of the results. The document would be clearer, however, if the details about exposure levels, purity of the compound, number of animals exposed, etc. were not repeated multiple times within the document. The reader gets bogged down by this repetition; one could, for example, present a single table for the oral studies and one for the inhalation studies that provides all of this information and then not have to repeat the description in multiple sections.

The data for genotoxicity and carcinogenic hazard are so scarce that answering this question for cancer is moot.

Mohammad I. Sabri

The draft document accurately summarized the background information and dose-response assessment for non-cancer and cancer hazards of 2-hexanone. The Review is concise and clear. However, the evolution of giant axonal swellings induced by 2,5-hexanedione (2,5-HD), the proximal metabolite of 2-hexanone, as the sensitive endpoint is not adequately considered and discussed.

Alan H. Stern

For the most part, the review document is clearly and concisely written. The review of the scientific literature is logical, objective and well written. The evidence for non-cancer hazard (as well as the slim evidence bearing on cancer hazard) is well synthesized. The section on the dose-response assessment, however, is less so. That section does not proceed in a logical stepwise manner and assumptions and choices were made that are not discussed or not fully discussed (see below).

Bernard Weiss

The document presents the available information in enough detail that I was able to follow the discussion, but I did not feel overwhelmed by particulars. It would have been useful, however, for the beginning of each section to present the reader with the eventual conclusion. For example, an introductory paragraph in Section 4.2 could have prepared the reader to understand why O'Donoghue (1978) would be the eventual choice as the critical experiment for oral exposure. It is only after slogging through a number of less relevant studies that we learn the eventual choice. Guiding the reader to a conclusion is the mark of a well-structured review.

2. Please identify any additional studies that should be considered in the assessment of the noncancer and cancer health effects of 2-hexanone.

Bahie Abou-Donia

The study of Abdo et al, 1982, used five concentrations (i.e., 10, 50, 100, 200, and 400 ppm) which exposed hens was not selected for the RfD assessment because "hens are not a suitable model for extrapolating experimental results to humans." This reason is not valid, because the EPA requires the adult hen to test organophosphorus for delayed neurotoxicity (OPIDN).

A. John Bailer

I don't know of any other studies that should be considered.

Frederick J. Miller

This reviewer is not aware of other published studies that should be considered in the assessment of potential noncancer or carcinogenic effects of 2-hexanone. However, the database is not large for this compound, and clearly, there can be identified areas that additional studies would be warranted.

Mohammad I. Sabri

To-date there is no scientific evidence that 2-hexanone causes cancer in animals or humans. Distal axonopathy is the most sensitive response and cardinal feature of 2-hexanone exposure. The following original papers may be considered in the assessment of 2-hexanone-neurotoxicity:

Spencer, P and Schaumburg, H (1977) Ultra structural studies in the dying-back process. III. The evolution of giant axonal degeneration. J. Neuropathol. Exp. Neurol. **36**:276.

Additional studies that should be considered are related to the coexposure of 2-hexanone or 2,5-HD with methyl ethyl ketone (MEK). Co-exposure with MEK produces nerve-fiber changes earlier and the toxic effects are more severe than those exposed to either solvent alone (Altenkirch, et al. (1982) J. Neurol. Sci, **57**:209; O'Donoghue, et al. (1984) Toxicol Appl. Pharmacol., **72**: 201).

Alan H. Stern

I am not aware of any additional studies that bear on the cancer or non-cancer health effects of 2-hexanone.

Bernard Weiss

Unfortunately, the literature provides rather limited choices. By establishing its neurotoxicity, O'Donoghue et al discouraged the further use of 2-hexanone as a solvent, which had the effect of dampening interest in it.

3. Please discuss research that you think would be likely to increase confidence in the database for future assessments of 2-hexanone.

Bahie Abou-Donia

A. Proposed studies

The following information either not available or are very limited. If they were available, the data base would have been stronger:

1. Immunotoxicity studies
2. Developmental Toxicity
3. Pharmacokinetics

B. The following references may be important, but are missing from the review. They came from my laboratory and not the result of a comprehensive search of the references. Other relevant references may also be missing.

Abou-Donia, M.B., D.M. Lapadula, G.M. Campbell, and P.R. Timmons (1985). The synergism of *n*-hexane-induced neurotoxicity by methyl *iso*-butyl ketone following subchronic (90 days) inhalation in hens: Induction of hepatic microsomal cytochrome P 450. *Toxicol. Appl. Pharmacol* **81:** 1-16.

Nomeir, A.A. and M.B. Abou-Donia (1985). Analysis of *n*-hexane, 2-hexanone, and 2,5 hexanedione, and related chemicals by capillary gas chromatography and high-pressure liquid cheomatography. *Anal. Biochem*. **151:** 381-388.

Lapadula, D.M., R.D. Irwin, E. Suwita, and M.B. Abou-Donia (1986). Crosslinking of neurofilament proteins of rat spinal cord *in vivo* after administration of 2,5-hexanedione. *J. Neurochem*. **46:** 1843-1850.

Lapadula, D.M., H.A. Tilson, G. Campbell and M.B. Abou-Donia (1987). Neurotoxic effects of combined treatment of 2,5-hexanedione and triethyllead chloride. *J. Toxicol. Environ. Health* **21:** 483-492.

Suwita, E., A.A. Nomeir and M.B. Abou-Donia (1987). Disposition, pharmacokinetics, and metabolism of a dermal dose of [^{14}C]2,5-hexanedione in hens. *Drug Metab. Disp.* **15:** 779-785.

Lapadula, D.M., Suwita, E., and Abou-Donia, M.B. (1988). Evidence for multiple mechanisms responsible for 2,5-hexanedione induced neuropathy. *Brain Res*. 123-131.

Lapadula, D.M., Habig, C., Gupta, R.P. and Abou-Donia, M.B. (1991). Induction of cytochrome P-450 isozymes by simultaneous inhalation exposure of hens to *n*-hexane and methyl *iso*-butyl ketone (M*i*BK). *Biochem. Pharmacol*. **41**: 877-883.

A. John Bailer

There are insufficient data to determine the carcinogenic potential of 2-hexanone. As such, data from a long-term carcinogenicity bioassay would be informative. Given that neurotoxicity is the primary outcome of concern, it is not clear that this would be useful for the overall assessment of 2-hexanone effects. The RfD and RfC are both based on relatively small studies with the RfC based on data "extracted" from figures in a paper and the variability calculated from values of a test statistic. The use of the actual data would clearly be preferred.

Frederick J. Miller

There is clearly a need to develop a dosimetry model (e.g., a physiologically-based pharmacokinetic model) for 2-hexanone. The limited available data infer that there are nonlinearities in absorption and transport mechanisms for this compound. The toxic agent is apparently one of the metabolites and there may well be species differences in how 2-hexanone is handled by the body. The ability to use the toxicity data to evaluate potential risk to humans from exposure would be greatly facilitated by dosimetry modeling.

The current document makes the point that there are no data on potential immunological effects in humans and that some white blood count changes in rat studies raise a cautionary flag. The magnitude of the changes seen in rats definitely warrants additional research in this area. While one could also conduct standard reproductive and developmental studies in animals, this reviewer would put a greater priority on additional studies to examine exposure level and duration relationships on the development of the neuropathies used to establish the RfC. Since the assumption in the current derivation of the RfC that the concentration time product is constant is clearly invalid (see response to Question C6), such studies would provide a better understanding of the concentration – duration of exposure response surface. This would then allow an RfC to be established that is based upon changes that precede overt toxicity and that prevent more serious effects on reproductive and developmental systems. The neurotoxicity studies should include functional and behavioral endpoints.

Mohammad I. Sabri

The following original research may increase the confidence in the database for future assessment of 2-hexanone. In vitro studies have shown that both 2-hexanone and 2,5-HD inhibit the enzymes of energy metabolism (Sabri et al., J. Neurochem. (1978) **32:** 683; Sabri, Brain Res. (1984) **297**: 145). Although, mM concentrations of 2-hexanone and 2,5-HD were used in these studies, the cumulative inhibitory effects of low doses of 2-hexanone and 2,5-HD should not be ignored. The inhibition of energy metabolism by 2-hexanone and 2,5-HD may disrupt energy-dependent axonal transport system and induce axonal degeneration. In vivo administration of 2-hexanone or 2,5-HD has also been shown to disrupt energy-dependent axonal transport system (Mendell et al.(1977) Brain Res., **133**: 107; Spencer and Griffin, In: Axoplasmic Transport in Physiology and Pathology, edited by D.Weiss and A. Gorio, Springer-Verlag, Berlin, page 92, 1982;

Breandgaard and Sidenius (1986) Brain Res., **378**: 1; Sickles (1989) Neurotoxicology, 10, 103). Some investigators have also reported acceleration of neurofilament transport in peripheral nerves and optic axons, Monaco et al. (1985) Proc. Natl. Acad. Sci., **82**: 920, and Monaco et al. (1989) Brain Res., **491**: 328). The following three review articles should be a good resource and likely increase the confidence in the database of 2-hexanone toxicity. (Spencer, et al. (1980) Crit. Rev. Toxicol, 7: 279; Couri and Milks (1982) Ann. Rev. Pharmacol. Toxicol. 22: 145; A. DeCaprio, In: Experimental and Clinical Neurotoxicology, eds. Peter Spencer and Herbert Schaumburg, Oxford University Press, page 633, 2000).

Alan H. Stern

Given the suggestion of the potential for effects of 2-hexanone based on the significant reduction in white cell count, as well as the contribution of this reduction to potential immunotoxicity, it would appear that immunotoxicity testing should be a first priority for additional research. This could also have the practical effect of reducing the UF for database uncertainty for both the RfD and RfC. A, 2-year chronic bioassays using pure 2-hexanone examining both non-cancer effects as well as neoplasias by the oral and inhalation routes would also be useful. In addition, a 2-generation combined reproductive-developmental study would also be very useful. Also useful would be studies elucidating the mechanism for the neurological effects of 2-hexanone. However, the absence of such studies should not delay the derivation of an RfD and RfC with the current information.

Bernard Weiss

A wider span of functional indices is required to better capture dose-response data. Neuropathology is the end result of a toxic process; functional measures would make it possible to trace its time course from the beginning. For example, O'Donoghue et al would have benefited by measures such as hindlimb foot splay, rotarod performance, and gait stability, all of which may reflect motor dysfunction and, in addition, provide continuous measures more suitable for quantitative risk assessment than ratings. In addition, the possibility of impaired somatosensory function also calls for appropriate testing. For inhalation, the choice of Johnson et al (1979) was appropriate. Although conduction velocity is a functional index, it does not, however, always parallel the results of behavioral testing. Adding motor function tests for both the rat and monkey subjects would offer a more consistent portrait of neurotoxicity.

4. Please comment on the identification and characterization of sources of uncertainty in sections 5 and 6 of the assessment document. Please comment on whether the key sources of uncertainty have been adequately discussed. Have the choices and assumptions made in the discussion of uncertainty been transparently and objectively described? Has the impact of the uncertainty on the assessment been transparently and objectively described?

Bahie Abou-Donia

The following uncertainty factors were used for RfD:

1. Intraspecies uncertainty factor (UFH), 10
2. Interspecies uncertainty factor (UFA), 10
3. Database deficiencies (UFD), 3; this should have been 10 because of the missing. studies on the immune system and developmental effects.
4. Total should be 1,000.

The following uncertainty factors were used for RfC:

1. Intraspecies uncertainty factor (UFH), 10.
2. Interspecies uncertainty factor (UFA), 10.
3. Database deficiencies (UFD), 3; this should have been 10 because of the missing. studies on the immune system and developmental effects.
4. NOAEL to LOAEL uncertainty factor, 3.
5. Total should be 3,000.

A. John Bailer

The RfD calculation included uncertainty factors for intraspecies uncertainty (10), interspecies (10), and database uncertainty (3) [see p. 61]. The RfC calculation included uncertainty factors for intraspecies uncertainty (10), subchronic-to-chronic (10), and database uncertainty (3) [see p. 73]. In addition, an adjustment for partition coefficient differences between monkeys and humans was used. Given the manipulations and variability assumptions required for the RfC calculation, one might argue for a database uncertainty factor > 3 for this calculation. The other uncertainty factors seem consistent with standard practice. The graphical displays (e.g. Fig 5-1) capture impact of the uncertainty adjustments on the BMD points of departure (or the NOAEL/LOAEL). Given the emphasis on BMD points of departure, it is unclear why NOAEL points of departure with uncertainty adjustments are displayed.

Frederick J. Miller

These sections of the document were well constructed and described. The uncertainties around the various studies and the rationale for which studies were chosen for the oral and inhalation reference analyses were clearly articulated. The assumptions that had to be

made were simply stated but not defended. For example, the assumption of a constant C x T product was invoked but has been shown in the literature to seldom be a valid assumption (see Question C6). There was no discussion about how this assumption is problematic for deriving an inhalation reference concentration. As such, the impact of this assumption was not adequately treated.

Mohammad I. Sabri

The choices and assumptions for the identification and characterization of sources of uncertainty factors (UF) are objectively described and discussed in the document. An intra-species uncertainty factor of 10 is applied to account for human-to-human variability. A default interspecies uncertainty factor of 10 is applied for extrapolation from animals to humans, because no suitable data on the oral toxicity of 2-hexanone in humans are available. An UF of 3 is applied to account for database deficiencies. This UF does not appear to be enough to account for database deficiencies. This factor should be 10. Thus application of a composite UF of 10 x 10 x 10 = 1000 appears appropriate for the calculation of the RfD for 2-hexanone.

Alan H. Stern

With one major exception (see below), the uncertainty factor (UF) adjustments adequately addressed the sources of uncertainty that they are intended to address. For the RfC, the most significant source of inadequately addressed uncertainty, however, lies in the dose-response modeling, including the use of benchmark dose modeling with these data. While benchmark dose modeling has the advantage over the NOAEL/LOAEL approach of utilizing all the data, the main focus in the derivation of RfDs and RfCs is in the fit of the model in the low-dose/low-response region. However, with only two dose levels (and a zero-dose control), the lowest of which, by the end of the study, is a LOAEL for a significant adverse effect, the fit of the model in the region of interest is highly uncertain.

This is all the more so since, given the paucity of data, many of the available models fit with approximately the same goodness of fit, but give BMDLs that differ by 34-36% . None of these models have an *a priori* claim to greater biological plausibility and the selection of the model is based solely on goodness of fit. Note that for the RfC, the BMCL is 243% of the (eventual) LOAEL (243 ppm vs. 100 ppm). In the case of the RfC, this uncertainty is exacerbated by the fact that for three of the four endpoints modeled (monkey sciatic-tibial nerve MCV, rat sciatic-tibial nerve and rat ulnar nerve MCV), the data chosen for modeling (6 months in monkeys, and 25 weeks in rats) are the last observations for which there are measurements of the effect at each of the two doses (this is not specifically stated in the text and is only clear from examining the BMD analysis data in the appendix). However, in each of these cases, at these time points, the effect at the low dose does not reach statistical significance compared to the controls. At later time points the effect at the low dose does reach statistical significance, however, for these time points, there are no data for the effect at the high dose. This effectively precludes BMD modeling at these time points. Thus, the data chosen for modeling in the

document are misleading with respect to the shape of the dose-response at the time when the full adverse effect emerges. It appears that it would be more appropriate to derive the RfC based on a LOAEL at 10 months in monkeys or 29 weeks in rats. The significant uncertainty inherent in the choices regarding the BMD modeling is not discussed.

Furthermore, the uncertainty inherent in the estimation of the variance in the RfC is not addressed. The standard deviation used in Appendix B2 is identical for all dose groups. This is implausible in reality and indicates that the value is merely a rough estimate. The uncertainty inherent in this estimate is not acknowledged and its implication is not discussed.

With respect to the UF adjustments, it is stated that "the toxicokinetic component [of uncertainty in animal-to-human extrapolation] is mostly addressed by the use of the determination of a human equivalent concentration as described by the in the RfC methodology…" However, this would only be true if the referenced HEC methodology were actually employed. The HEC calculation, however, was not carried out because there were no data for the blood-gas partition coefficient for either monkeys or rats and a default value of one was used for ratio of animal to human blood-gas partition. Since the HEC methodology was not used, but rather, only considered, the toxicokinetic uncertainty in the animal-human extrapolation remains unaddressed. This is not discussed.

Bernard Weiss

There is not much to add. The review's limitations stem from the lack of information in the literature. I would not rate the adequacy of the database, therefore, as "medium." I would confer a rating between medium and inadequate, partly because of the endpoints chosen. For oral exposure, the criterion was neuropathology, while, as noted above, functional indices might have revealed effects at lower levels. For the inhalation studies, MCV seems to have provided the most transparent criterion, so I would rate database adequacy as "medium."

(B) Oral reference dose (RfD) for 2-hexanone

1. A chronic RfD for 2-hexanone has been derived from a 13-month drinking water study (O'Donoghue et al., 1978) in male rats. Please comment on whether the selection of this study as the principal study has been scientifically justified. Has this study been transparently and objectively described in the document? Please identify and provide the rationale for any other studies that should be selected as the principal study.

Bahie Abou-Donia

The EPA report included justification for the selection of 13-month drinking water study (O'Donaghue et al, 1978) that seems convincing.

A. John Bailer

This study exhibited a response that was argued as being most relevant for a potential neurotoxin. In addition, these data exhibited a strong dose-response relationship even with relatively few animals tested at each treatment (dose).

Frederick J. Miller

The selection of the O'Donoghue et al. (19780 study in male rats was as the study for deriving a chronic RfD for 2-hexanone was scientifically justified within the document. The description of the study was clear and the basis for selecting it compared with other available studies was objectively presented.

Mohammad I. Sabri

To date the 13-month chronic drinking water study (O'Donoghue et al., 1978) in rats appears to be the most extensive and suitable principal study for the derivation of the RfD of 2-hexanone. The experimental details of the study have been objectively described and discussed with a few exceptions. For example the study has documented the intake of 2-hexanone and reduction in body weight, but food and water intake data are not reported. Although, five sub-chronic studies have been cited as the supporting studies, they do not qualify as the principal studies.

Alan H. Stern

The choice of O'Donoghue et al. (1978) as the principal study appears reasonable and has been reasonably justified in the document. In particular, it is the only chronic study, and the only study to use multiple dose levels. The doses are reasonably spaced and amenable to benchmark dose modeling. The text mentions that the Abdel-Rahman et al. (1978) study, although utilizing lower doses, only reported data for the first 4 weeks of the study. Having read that paper, it should also be noted that data were not reported for the 0.1% exposure level (i.e., the low dose group). The O'Donoghue et al. (1978) paper,

itself provides complete data that are clearly presented. The review document provides a good and objective summary of those data.

Bernard Weiss

I agree with the authors and previous reviewers that this study provides the most useful source of data for calculating risk assessment parameters. It was carefully conducted in a highly-regarded laboratory and comes quite close, actually, to the GLP standards adopted subsequently.

2. Myofibrillar atrophy of the quadriceps muscle was selected as the critical effect. Please comment on whether the rationale for the selection of myofibrillar atrophy as the critical effect has been scientifically justified. Has this selection been transparently and objectively described in the document? Please provide detailed explanation. Please comment on the selection of myofibrillar atrophy of the quadriceps muscle as the critical effect rather than other endpoints identified in O'Donoghue et al. (1978). Please comment on the selection of myofibrillar atrophy of the quadriceps muscle as compared to the peripheral nerve axonal swelling. Please identify and provide the rationale for any other endpoints that should be considered in the selection of the critical effect.

Bahie Abou-Donia

Despite the EPA's justification for the selection of myofibrillar atrophy of the quadriceps muscle as the critical effect, the use of giant axonal swelling and degeneration in the central and peripheral nervous systems might have been a more appropriate choice because these alterations are characteristics of 2-hexanone-induced neuropathology that accompanies its neurotoxicity.

A. John Bailer

In terms of the ultimate BMDL derived from these two responses (49.9 for the quadriceps and 69. 2 for the calf muscle), the selection of the site did not matter much in terms of the point of departure for the RfD calculation. This is a small difference relative to the uncertainty factors that are then applied. I defer to my toxicology and physiology colleagues to comment on the relevance of these endpoints.

Frederick J. Miller

The rationale for selecting myofibrillar atrophy of the quadriceps muscle and the calf muscle as critical effects was adequately defended. The document noted that these endpoints were chosen over other neuropathologic endpoints because they occur due to axonal atrophy, which an endpoint is shown to be best correlated with nerve dysfunction independent of the route of exposure. However, discussions at the 2-hexanone peer review meeting challenged the dismissal of axonal swelling in light of the fact that this occurs by filament cross linking that is irreversible. Moreover, neurotoxicologists on the panel indicated that this effect accumulates as long as exposure continues.

The selected endpoints, chosen by the Agency for BMD analysis, had clear dose-dependent responses that were amenable to BMD modeling. In addition, these endpoints represent adverse changes compared to changes such as peripheral nerve axonal swelling. As noted in the document, axonal swelling does not necessarily lead to nerve dysfunction. This aspect combined with the fact that the axonal swelling data: (1) did not show dose-dependent changes for the spinal cord, (2) showed only swelling for the dorsal root ganglia at the highest dose, and (3) showed marked dose-dependent changes for the brain and peripheral nerve led the authors to focus on endpoints showing

myofibrillar atrophy. Myofibrillar atrophy of the quadriceps muscle was chosen as the critical effect over myofibrillar atrophy of the calf muscle because it yielded a slightly lower BMDL. However, this reviewer would still have liked to see the BMD results for the brain and peripheral nerve as well.

Mohammad I. Sabri

Myofibrillar atrophy of the quadriceps muscle was chosen over a more sensitive endpoint (axonal swellings) of 2-hexanone neurotoxicity. Whereas axonal swellings were seen at all the dosing regimen of 2-hexanone, axonal atrophy was not even assessed in the principal study (O'Donoghue et al.1978). 2-Hexanone was found to be quite neurotoxic at the doses used in the principal study (O'Donoghue, et al., 1978) so much so that even the lowest dose (143 mg/kg/day) produced axonal swellings in 8/10 animals, as opposed to myofibrillar atrophy in quadriceps muscle where only 1/10 animals showed muscle atrophy. Clearly the presence of axonal swellings, not muscle atrophy, is a more sensitive endpoint of 2-hexanone-toxicity. Dosing the animals with lower concentrations of 2-hexanone would be expected to induce axonal swellings in a dose-response manner. Although, the incidences of myofibrillar atrophy of the quadriceps muscle and the calf muscle displayed a dose-dependent response, muscle atrophy was seen at higher doses relative to axonal swellings. Axonal swellings precede axonal degeneration, muscle atrophy and limb paralysis regardless of the route of 2-hexanone administration. Single teased fibers studies have also shown the presence of giant axonal swellings and distal nerve-fiber atrophy in 2,5-HD neuropathy (Spence and Schaumburg, 1977). However, some investigators reported that fiber atrophy, not giant axonal swelling, is a primary neuropathologic feature of gamma-diketone neuropathy.

Alan H. Stern

In the O'Donoghue et al. (1978) study, several statistically significant adverse effects were noted: decreased body weight; increased relative liver, kidney and testes weight; and axonal effects (swelling and atrophy). It should be noted that in the description of this study (pg. 24), the document states that, "Other than neural effects and changes in body weight, no clinical signs… were found." It is not clear why the effects on organ weight were excluded from this summary statement. Giant axon (quadriceps and calf) atrophy, ultimately identified as the critical effect, is clearly an appropriate effect for development of an RfD. It is sub-clinical, but clearly adverse and predictive of clinical effects. However, it is not clear that among the adverse effects noted in that study, it is the most sensitive (i.e., would produce the lowest RfD). In this study, a decrease of 10% in body weight was observed between the lowest concentration in drinking water (0.25%) and the middle concentration (0.5%). Decrease of greater than 10% in body weight is a relatively common endpoint in IRIS for derivation of RfDs. Decreased body weight is a less specific endpoint than axonal atrophy and can result from non-toxicological factors (e.g., decreased water consumption due to potability issues). Nonetheless, in the absence of obvious non-toxicological factors, body weight decrease is a valid and appropriate endpoint for RfD derivation. It is not necessary that it be in the same chain of causality as a more specific effect (i.e., axonal effects). Section 5.1.1, that discusses the choice of

the critical effect notes the presence of decreased body weight, but does not discuss why it was not chosen Changes in organ weight, particularly increases, can be adaptive rather than strictly adverse, however, no such case is made, and changes in organ weight were not even mentioned in section 5.1.1.

The underlying rationale for the choice of axonal atrophy appears to be the specificity and uniqueness of this effect to 2-hexanone and related compounds metabolized to 2,5-hexanedione. However, under the definition of the RfD, neither the degree of adversity, nor the specificity, nor the uniqueness of an endpoint need drive the RfD selection if the endpoint that is ultimately selected is protective of that endpoint even if it is not causally or mechanistically related to the endpoint. Given this reasoning, body weight decrease (and possibly organ weight increases- depending on their toxicological interpretation) should have been examined with benchmark dose modeling and its point-of-departure should have been compared to that generated for axonal atrophy.

The rationale for selection of axonal atrophy as opposed to axonal swelling – that axonal swelling does not correlate well with clinical effects and does not necessarily progress – is contradicted by information provided by Drs. Abou-Donia and Sabri. They indicate that axonal swelling occurs "upstream" of axonal atrophy and is progressive and predictive of axonal atrophy with continued exposure at the same dose. Given this, EPA needs to explain why axonal swelling rather than axonal atrophy should not be identified as the critical effect.

Bernard Weiss

Table 4-8 provides no clear rationale, in itself, for choosing quadriceps myofibrillar atrophy as the critical effect. On the basis of dose-response functions, brain axonal swelling or calf muscle atrophy could have served as well and would have yielded, as later calculated, about the same RfD. The choice is well-argued, however, and based on consistency, but, so to speak, quantitatively uncomfortable. It also struck me that body weight reduction and liver weight increase were also dose-related. These effects deserve further comment.

3. Please comment on the selection of the uncertainty factors applied to the point of departure (POD) for the derivation of the RfDs. For instance, are they scientifically justified? Are they transparently and objectively described in the document?

Bahie Abou-Donia

The selection of the uncertainty factors applied to the point of departure (POD) for the derivation of the RfDs was justified in the report.

A. John Bailer

As noted above, the RfD calculation included uncertainty factors for intraspecies uncertainty (10), interspecies (10), and database uncertainty (3) [see p. 61]. These seemed reasonable given standard practice for RfD calculations.

Frederick J. Miller

The material on page 61 of the document covers the selection of the values of various uncertainty factors as well as the rationale for their selection. The authors adequately covered this topic.

Mohammad I. Sabri

The selection of the uncertainty factors of a 10% extra risk (ER) of myofibrillar atrophy of quadriceps muscle and calf muscle for the point of departure in the derivation of the RfD appears appropriate, and justified. It represents a minimal biologically significant change. These points are objectively discussed and described in the document.

Alan H. Stern

The choice of uncertainty factors (UFs) is clearly and objectively described. The choice of a UF of 10 for sensitive humans, and animal-to-human extrapolation is clear and justified. The only question arises in considering the assignment of a UF of 3 for database insufficiency. This is addressed in the following charge question.

Bernard Weiss

The UFs comply with EPA practice and are clearly explained. I agree with the 10% figure.

4. Please comment specifically on the database uncertainty factor of 3 applied in the RfD derivation. Please comment on body of information regarding reproductive, developmental toxicity, and immunotoxicity on 2-hexanone as well as the relevance of toxicity data on n-hexane in the determination of the database uncertainty factor. Please comment on whether the selection of the database uncertainty factor for the RfD has been scientifically justified. Has this selection been transparently and objectively described in the document?

Bahie Abou-Donia

The information regarding developmental toxicity and immunotoxicity is inadequate or missing. The database uncertainty factor should be 10 instead of 3.

A. John Bailer

I don't have a strong opinion about this.

Frederick J. Miller

While the UF of 3 for the database initially appeared reasonable to this reviewer, discussions at the peer review meeting changed my viewpoint such that I think an UF of 10 is more appropriate. The fact that extremely high doses have been used in reproductive studies with n-hexane, which is a precursor of 2-hexanone, and that these studies are negative does not get 2-hexanone "off the hook" because this conversion is very slow and not likely to shed any light on any effects that 2-hexanone may impart on these systems. The authors cite the suggestion from inhalation studies of possible concern for immunotoxicity and reproductive toxicity as the overall basis for selecting a UF of 3 for the quality of the database. This does not really stand the test of adequacy support, and, in view of the paucity of database, an UF of 10 would be appropriate.

Mohammad I. Sabri

An uncertainty factor of 3 was applied to account for database deficiencies in the derivation of the RfD for 2-hexanone. This factor does not account appropriately for reproductive, developmental and immunotoxic effects of 2-hexanone. Some information on the developmental studies with *n*-hexane, a precursor of 2-hexanone, and 2,5-HD, the metabolite of 2-hexanone, are available however, studies on the toxic effects of oral 2-hexanone exposure are not available. Thus an uncertainty factor of 3 is not justified. An UF of 10 is suggested.

Alan H. Stern

While the lack of developmental and multigenerational reproductive studies is ameliorated somewhat by data from studies of n-hexane and 2,5-hexanedione, those results are not unequivocally negative. It is noteworthy that the developmental effects

seen in rats from maternal injection of 2,5-hexanedione are similar to those seen for adult exposure to 2-hexanone. The immunotoxic effects seen with inhalation exposure are, however, cited in the text as the primary reason for the UF of 3 (as opposed to no UF for database insufficiency). Apparently, the value of 3 rather than 10 was chosen because those observations are from inhalation studies and no similar data exist for ingestion studies. If the rationale for not assigning a full UF is, indeed, that the available data do not relate directly to the route of exposure under consideration, and if this is considered a potentially significant endpoint (this is not made clear in the text) then, then the lack of ingestion specific data should rather justify a full UF of 10 rather than be interpreted as an indication of reduced uncertainty. This needs to be clarified in the text.

Bernard Weiss

An additional UF is warranted. The database on which the RfD depends is fragmented and, in many respects, is rather crude by criteria such as sensitive and appropriate endpoints, such as functional measures. Because of the absence of clear information about immunotoxicity and developmental neurotoxicity, the additional UF should be a factor of 10.

5. Please provide any other comments on the derivation of the RfD.

Bahie Abou-Donia

The total uncertainty factors should be 1000 instead of 300.

A. John Bailer

As a noted in the specific comments below, a 4 parameter model is fit to 4 data points [B-3]. It is no surprise that this generates a good fit. In addition, the parameter convergence criteria was set to 10^(-8) and an estimate of beta3 was 3.7x10^(-8) was obtained. Do any concerns about convergence exist? Should you redo this with tighter convergence criteria? One option would be to rescale the doses and then refit the models. The parameter estimates would no longer be so close to the convergence criterion. As an aside, the BMDL varied by a factor of 3.5 over all of the models fit. This model uncertainty is a small factor relative to the other uncertainties considered for the RfD.

Frederick J. Miller

The fact that the critical study selected involved continuous exposure saved the Agency from having to invoke a constant product of concentration and time for level of effect. As discussed in Question C6, such an assumption is seldom valid.

Mohammad I. Sabri

Although five sub-chronic studies have been published on the toxicity of 2-hexanone, the chronic oral report of O'Donoghue, et al., 1978, appears justified as the principal study for the derivation of RfD.

Alan H. Stern

The version of the BMDS software used in the document is version 1.3.2. This is an older version. Version 1.4.1c has been available for some time. While it is understandable that this draft has been in preparation and review for some time, the calculations should reference the revised version of this software. I have used the newer version of the software to calculate the BMDLs. The following tables compare the values reported in the review document with the values I obtained using the same input parameters in version 1.4.1c. Note that the quantal quadratic model is identified in the document as yielding the most appropriate BMDL for myofibrillar atrophy of the calf muscle. This model is not available in the later version of the BMDS software. It is not clear why that model was eliminated. However, that should be addressed. Also, note that the newer version of the software does not calculate the AIC statistic given with the results for the older model. Rather, the newer version provides the chi^2 statistic.

Quadriceps muscle atrophy

model	**AIC v.1.3.2**	**p-value**	**BMDL v.1.3.2**	**Chi^2 v.1.4.1c**	**p-value**	**BMDL v.1.4.1c**
gamma	24.6565	0.9034	86.2822	0.2	0.9034	86.2822
log-logistic	25.1541	0.7602	96.3847	0.55	0.7602	98.3847
logistic	24.5648	0.9392	94.5301	0.13	0,9392	94.5301
multistage	**22.3952**	0.9995	49.9434	0.02	0.9995	49.9434
log probit	24.9892	0.7974	98.9646	0.45	0.7974	98.9646
probit	24.4241	0.9822	86.9997	0.04	0.9822	86.9997
quantal linear	29.9156	0.1628	22.8722	defaults to weibull		
quantal quadratic	23.7952	0.8118	78.3857	not available		
weibull	24.3816	0.9945	78.2565	0.01	0.9945	78.2565

For the quadriceps muscle myofibrillar atrophy, the BMDLs for models that are present in both versions are identical. However, note that when evaluating the model fit based on the chi^2 statistic, the Weibull and multistage models give almost identical fits, more so than when comparing the fits using the AIC statistic. It is not clear what the practical significance of this difference is. The multistage model gives a significantly smaller BMDL and this may be the rationale for selecting that model over the Weibull. However, if that is the case, it should be explicitly addressed. Alternatively, in the absence of a specific biological rationale for selection of one model over another, EPA should give careful consideration to averaging (or applying a weighted average to) the BMDLs that arise from models with essentially identical fits.

Calf muscle atrophy

model	**AIC v.1.3.2**	**p-value**	**BMDL v.1.3.2**	**Chi^2 v.1.4.1c**	**p-value**	**BMDL v.1.4.1c**
gamma	27.7837	0.8972	48.9374	0.22	0.8972	48.9374
log-logistic	28.36	0.7402	63.2262	0.60	0.7402	63.2262
logistic	27.9769	0.8457	70.8106	0.34	0.8457	70.8106
multistage	**27.4841**	**0.9956**	**30.1238**	**0.10**	**0.9500**	**35.0954**
log probit	28.0906	0.8019	67.2193	0.60	0.8019	67.2193
probit	27.7116	0.9227	65.7944	0.16	0.9227	65.7944
quantal linear	30.2036	0.3498	19.0312	defaults to weibull		
quantal quadratic	25.8664	0.9701	69.2097	not available		
weibull	27.5386	0.9756	45.8927	0.05	0.9756	45.8927

For the calf muscle myofibrillar atrophy, with the exception of the multistage model, the models that are present in both software versions give identical BMDLs. However, with the absence of the quantal quadratic model in the newer software, the best fit is given by the Weibull model. The multistage model gives a very similar fit. However, note that the BMDL calculated with the newer software is not identical with the BMDL reported in the document using the older software. This is significant given the comparability of the fits of the multistage and Weibull models and the observation that the Weibull model gives the marginally better fit, but the multistage model gives a significantly lower BMDL. This should he addressed.

Bernard Weiss

Quantification of endpoint is critical in elaborating an RfD. In this instance, we are contending with unquantified ratings, that is, a rough ordinal scale. Quantitative morphometry has yet to be appled to peripheral nerve damage, which is why the use of MCV, as in the inhalation study by Johnson et al (1979) represents an advantage. Given the many measures available for assessing motor function in rodents, any future studies with n-hexanone should take advantage of these possibilities.

(C) Inhalation reference concentration (RfC) for 2-hexanone

1. A chronic RfC for 2-hexanone has been derived from a 10-month inhalation study (Johnson et al., 1977) in rats and monkeys. Please comment on whether the selection of this study as the principal study has been scientifically justified. Has this study been transparently and objectively described in the document? Please comment on the use of a 10-month monkey study (Johnson et al., 1977) as opposed to a 72-week rat study (Krasavage and O'Donoghue, 1977). Please identify and provide the rationale for any other studies that should be selected as the principal study.

Bahie Abou-Donia

The 10-month inhalation study (Johnson et al, 1977) in rats and monkeys was adequately justified by the EPA. The study was objectively described in the document. As an inhalation study, it was more appropriate to be used in deriving the RfC than Krasavage and O'Donoghue, 1977.

A. John Bailer

This appears to be one of the few studies with "strong" dose-response information which facilitates the construction of BMDs/BMDLs (only 3 doses tested: 0, 100, 1000 ppm). The report opted to focus on a particular exposure duration, and needed to extract dose-response information from a figure and to construct variability estimates from a test statistic and an assumption of equal variability. Access to original, raw data (or even published summary statistics) would be preferred. The use of a non-human primate gives greater confidence for species extrapolation. As an aside, neither study (as displayed in Table 5-3) exhibited strong low-dose effects. Finally, the rejection of the Krasavage and O'Donoghue (1977) study results appeared to be justified based on "purity of 2-hexanone … not stated" and "difficult to ascertain if MiBK impacted the toxicity of 2-hexanone" [p. 65].

Frederick J. Miller

The authors have scientifically justified the selection of the inhalation study by Johnson et al. (1977) as being the principal study to use in deriving an RfC; moreover, the salient features of the study were well presented. Having a study in primates is definitely a plus since their respiratory tract geometry is quite similar to humans. There was strong similarity of the results in both primates and rats. The authors make it clear that, while the rat study of Krasavage and O'Donoghue was longer in duration, there was insufficient data presented in the study to derive a reliable RfC. In addition, the study in monkeys was sufficiently long to represent chronic exposure and had the added feature of being in a species more similar to humans.

There are no other studies discussed in the document or that this reviewer is aware of that would represent a better choice for the principal study to use in deriving an RfC for 2-hexanone.

Mohammad I. Sabri

The use of a 10-month inhalation study (Johnson et al., 1977) for the derivation of the RfC of 2-hexanone appears justified as this study was performed in two different animal species- monkeys and rats- exposed to two different concentrations 6-hour per day 5day per week. Appropriate control groups were exposed to zero concentration of 2-hexanone. The study appears to be objectively described and discussed in the document. As monkeys and humans have a similar respiratory tract and breathing patterns are similar, the 10-month study of Johnson et al., 1977, appears appropriate for the RfC calculation as opposed to a 72-week rat study of Krasavage and O'Donoghue, 1977.

Alan H. Stern

The basis for the selection of the Johnson et al. study was clearly presented. The charge question's focus on Krasavage and O'Donoghue (1977) seems odd since Krasavage and O'Donoghue (1979) is a longer study that is described in greater detail in the document. . The statement in section 5.2.1 in the discussion of the choice of the principal study that, "Also, Krasavage and O'Donoghue (1979) provide limited information to serve as the basis for a reference concentration" does not appear to have been dealt with or described in further detail elsewhere in the text and requires further explanation. Nonetheless, in both of the Krasavage and O'Donoghue studies, the doses were the same as those in Johnson et al (1977). The lowest dose, however, was a NOAEL in the Krasavage and O'Donoghue studies, while it was a LOAEL in the Johnson et al. study. Thus, the longer duration of the Krasavage and O'Donoghue studies notwithstanding, the observation of adverse effects at the lower dose in the Johnson et al. (1977) study, warrants its selection as the principal study.

It should be noted, however, that the description of the Johnson et al. study as a "10 month" study is misleading because only the 6-month data were used for the sciatic-tibial nerve effect in monkeys. The data for the remaining 4 months were not used even though they yield a LOAEL for the low dose as opposed to the 6-month data that yield a LOAEL for the high dose.

Bernard Weiss

Johnson et al (1977) was the proper choice, principally because of their choice of a functional endpoint. Krasavage and O'Donoghue (1977) and Egan et al (1980) disagreed about the effects of 100 ppm, perhaps because of the low reproducibility of neuropathological ratings. In addition, inclusion of a nonhuman primate cohort diminishes the problem of extrapolation from exclusively rodent species to humans.

2. Motor conduction velocity of the sciatic-tibial nerve in monkeys was selected as the critical toxicological effect. Please comment on whether the selection of this critical effect has been scientifically justified. Has this selection been transparently and objectively described in the document? Please provide detailed explanation. Please comment on the use of motor conduction velocity of the sciatic-tibial nerve instead of motor conduction velocity of the ulnar nerve. Please comment on the use of monkey data instead of rat data. Please identify and provide the rationale for any other endpoints that should be considered in the selection of the critical effect.

Bahie Abou-Donia

Selection of the motor conduction velocity of the sciatic-tibial nerve in monkeys as the critical toxicological effect of 2-hexanone is appropriate because Sciatic and tibial nerves are prime targets for its toxicity. The report has scientifically justified this selection. The monkey data were justifiably used over those of the rat because the monkey is a primate.

A. John Bailer

I defer to my toxicology and physiology colleagues to comment on the relevance of these endpoints. The data in Table 5-3 might be better displayed in a figure (or figures) where the mean response is plotted versus exposure duration (months) with separate profiles for the 3 concentration levels of 2-hexanone (0, 100, 1000 ppm). As an aside, a figure might be provided for the data in Table 5-1. The incidence of myofibrillar atrophy of quadriceps or calf could be plotted vs. dose of 2-hexanone in drinking water (0, 143, 266, 560 mg/kg/d). The incidence of axonal swelling could be displayed in a separate figure or superimposed on the same figure.

Frederick J. Miller

The reasoning for selecting motor conduction velocity (MCV) of the sciatic-tibial nerve as the critical effect is presented on pages 65 and 66 of the document. Unfortunately, these pages do not bring forward a number of points concerning the effect levels and patterns that are described on pages 38 and 39, so the reasoning and defensibility of the selection of the critical endpoint and the use of the monkey data over the rat data are not well defended. For example, not discussed was the fact that effects were so pronounced (hind-limb drag) that cessation of exposure occurred after 25 weeks for all rats and monkeys exposed to 1000 ppm 2-hexanone. Also, monkeys in the low exposure group (100 ppm 2-hexanone) were exposed for 41 weeks while rats in the low exposure group were only exposed for a total of 29 weeks.

Effects on the sciatic-tibial nerve were seen at both 2-hexanone exposure levels in monkeys, with sporadic temporal effects seen in rats at the high exposure level. Importantly, the authors point out that the likely toxic moiety (2,5-hexanedione, a metabolite of 2-hexanone) usually affects long axons, such as the sciatic tibial nerve,

prior to other nerves. This adds strength to the reasoning behind selecting sciatic-tibial MCV as the primary effect.

The ulnar nerve pattern of effects was similar to that observed for the sciatic-tibial nerve. However the pattern and significance of the effects were more reproducible in rats compared to monkeys. The same could be said for MAP amplitude in response to sciatic stimulation.

RfC values were derived for both rats and monkeys for sciatic-nerve MCV and ulnar MCV. The monkey sciatic-nerve MCV appears to have been picked for three reasons: (1) monkeys are more similar to humans than are rats, and (2) the derived RfC for the sciatic-nerve MCV is slightly lower than that derived for the ulnar-nerve, and (30 the derived RfCs in both species for the sciatic-nerve MCV are lower than those for the ulnar-nerve MCV. This reasoning should be clearly articulated in Section 5.2.1.

Mohammad I. Sabri

The selection of the sciatic nerve MCV in monkeys is scientifically justified because biologically monkeys are more similar to humans than rats. This point is objectively described and discussed in the document. Although, 2-hexanone produces similar patterns of neuropathy in sciatic-tibial nerve and in the ulnar nerve, the sciatic-tibial nerve MCV was used for the derivation of the RfC because monkeys showed significant decrements in MCVs in sciatic-tibial nerve at the lowest concentration of 2-hexanone after 9 months of exposure. By contrast, ulnar nerve MCV initially showed significant decrease but after 6 months this decrease was not different from the control nerve.

Alan H. Stern

Although I do not have specific expertise in this area, it appears that decrease in MCV is a significant adverse effect. Nonetheless, it seems that this case is not made explicitly in the document. There should be a brief discussion of the relationship between decreased MCV and functional effects. With respect to the choice of the sciatic-tibial nerve versus the ulnar nerve, there is a brief mention in section 6.2.2 (pg. 80) that the sciatic-tibial nerve is chosen over the ulnar nerve because the former is longer and, therefore, more likely to experience degenerative effects earlier. However, this needs more discussion earlier in the text. For example, how much longer is the one than the other? What is the basis for this statement? Nonetheless, given that, for either monkeys or rats, the sciatic-tibial nerve yields the lower BMDL, it is an appropriate choice. With respect to the choice of monkeys versus rats, given that the BMCLs for monkeys and rats are quite similar, it is reasonable to select the monkey endpoint due to its greater relevance to humans even though the rat BMCL is slightly smaller. However, the RfCs do not appear to reflect the same relative proportions as the BMCLs. It is not clear why that is the case.

Bernard Weiss

I was not convinced that the sciatic-tibial nerve was a better choice than the ulnar nerve. The basis for the choice was statistical significance, but I would have liked to see the actual data. Statistical significance is not equivalent to functional significance, moreover. Ulnar nerve damage in humans is probably more disabling, at least subjectively, than sciatic-tibial nerve damage because it underlies more aspects of the kind of fine motor control we need in so many common tasks. If further research is ever conducted with nonhuman primates, this latter aspect could be assessed behaviorally as well.

3. Estimates of the standard deviation of the responses in each dose group are needed to calculate benchmark doses (BMDs) and their corresponding lower confidence limits (BMDLs). This information was not provided in Johnson et al. (1977), the principal study. Therefore, an indirect method for estimating this missing information on response variability was devised. Please comment on the procedure used to determine the standard deviation. Please comment on the use of digitization as a method to abstract data from Johnson et al. (1977) for the derivation of the inhalation reference concentration.

Bahie Abou-Donia

The report has adequately described and justified the indirect method used for estimating the missing information and the standard deviation, both of which were missing in the Johnson et al, 1977 study.

A. John Bailer

The strategy of working from the F statistic back to the pooled variance estimate was reasonable ASSUMING that the variance is common in the different treatment groups. This is a strong assumption for many studies where variability tends to increase with mean responses. Are you confident that equal variability in equal treatment group is a reasonable assumption for this response? The "digitization" of the data is unavoidable given this appears to the best dose-response study although the method of obtaining the mean data is not optimal.

Frederick J. Miller

The method used to estimate response variability given that such data were not reported in the published study is a statistically valid approach. The results from averaging the resulting variances should probably be compared to those obtained only using the specific time variance to see if this makes a difference. This reviewer doubts that the original authors analyzed the data in a multivariate repeated measures framework, which would have been the correct statistical analysis in order to account for the correlation inherent in repeated observations from the same monkeys.

The digitization procedure that was used for data capture from published figures was reasonable and did not introduce errors of any magnitude that would affect the analytical results.

Mohammad I. Sabri

The design and data of the study by Johnson et al. may have been acceptable in 1977, but may not be satisfactory today. A study done today would have expected statistically analyzable data for the calculation of benchmark concentrations and their corresponding lower confidence limits. An indirect method for estimating variability in MCV measurements was employed statistical tests using ANOVA and Student's *t*-test to

determine significant differences in mean MCVs across the exposure groups. Since no previous RfC assessment data exists on IRIS, data extraction via digitization may be an acceptable alternative for the derivation of the RfC.

Alan H. Stern

I am not qualified to comment in detail on the procedures used to estimate the standard deviation in each group from the data provided in Johnson et al. However, based on discussion among the reviewers, it is clear that the critical factor in this procedure is the assumption of uniform variance among the outcomes at the different doses. This assumption should be acknowledged and discussed.

Bernard Weiss

The methods used here were warranted.

4. Please comment on the selection of the uncertainty factors applied to the POD for the derivation of the RfCs. Are they scientifically justified? Are they transparently and objectively described in the document?

Bahie Abou-Donia

The selection of the uncertainty factors applied to the POD for the derivation of the RfCs was adequately justified and objectively described in the document. Because the dats base is not complete, the UF_D should be 10.

A. John Bailer

As I mentioned above, given the manipulations and variability assumptions required for the RfC calculation, one might argue for a database uncertainty factor > 3 for this calculation. The set of uncertainty factors is reasonably described.

Frederick J. Miller

The reasoning for the intraspecies default UF of 10 is readily apparent and defensible. This reviewer originally questions the need for an UF of 10 for a default of subchronic-to-chronic exposure rather than a value of 3 because the exposures were long enough that periodicity of the blood plasma and serum levels was achieved, thereby providing stable body burdens. For non-cancer endpoints and for exposures that do not overload lung clearance capabilities, this usually results in "smoldering lesions" that do not progress further. However, as brought out by discussions at the peer review meeting, this is not the case for 2-hexanone effects on nerves because filament cross linking is irreversible as well as cumulative. Hence, this reviewer endorses an UF of 10 for less than chronic exposures having been used.

Mohammad I. Sabri

The selection of the uncertainty factors (UF) applied to the point of departure (POD) to calculate the RfC is fully described, discussed and appears justified. A composite UF of 1000 was applied in the derivation of the RfC to allow for potentially sensitive human subpopulations, to account for the less-than-lifetime exposure, to adjust for uncertainties in extrapolating from rats to humans, and to account for database deficiencies. Although, the UF used for the POD in calculating the RfC is objectively described in the document, the overall confidence in the derivation of the RfC is medium.

Alan H. Stern

The narrative in section 5.2.4 is extremely confusing with respect to the species chosen for the RfC derivation. The third bullet, dealing with the UF for animal-to-human extrapolation, refers only to rats even though the beginning of this section clearly states that the monkey data were chosen for derivation of the RfC. Thus, it is not clear how the

monkey data were treated with respect to this UF. This may account for the change in the relative proportions of the rat and monkey UFs referred to above. With respect to the animal-to-human UF, the text states that, "…the toxicokinetic component is mostly addressed by the determination of a human equivalent concentration as described in the RfC methodology…." This is true only if the HEC methodology is actually employed. In HEC calculation, however, there were no data for the blood-gas partition coefficient for the monkeys and the default of 1.0 was used. Therefore, it appears that toxicokinetic uncertainty was *not* addressed, contrary to the text.

Bernard Weiss

The interspecies and individual variation UFs are standard. I liked the charts in Figure 5-2.

5. Please comment specifically on the database uncertainty factor of 3 applied in the RfC derivation. Please comment on body of information regarding reproductive, developmental toxicity (including developmental neurotoxicity), and immunotoxicity on 2-hexanone, as well as the comparability and relevance of toxicity data on n-hexane and 2,5-hexanedione in the determination of the database uncertainty factor. Please comment on whether the selection of the database uncertainty factor for the RfC has been scientifically justified. Has the selection of the database uncertainty factor been transparently and objectively described in the document?

Bahie Abou-Donia

The available literature regarding developmental toxicity and immunotoxicity is inadequate or missing. The database uncertainty factor should be 10 instead of 3.

A. John Bailer

See my response to the previous question.

Frederick J. Miller

Unclear to this reviewer was why the authors cited the use of an UF of 3 when extrapolating from rats to humans when in this instance they are going from monkeys to humans. However, at the peer review meeting, Agency staff clarified that this statement was actually intended to refer to "from monkeys to humans". The document would be clearer if separate bullets were used for discussing the animal to human UF compared with a UF for the strength and quality of the database.

Concerns about the lack of immunotoxicity, reproductive, and developmental toxicity studies are raised, but these concerns are folded into selection of a UF of 3 for the database. The rationale presented on page 74 that identifies 3 factors that led to choosing a UF of 3 for the strength and quality of the database is straight forward, but was, as discussion at the peer review meeting pointed out, not totally defensible. For example, the conversion of n-hexane to 2-hexanone is very slow and, so studies using n-hexane are not likely to be relevant to the toxicity of 2-hexanone. In light of this and in view of the fact that the Agency is using 6-month exposure results for BMD modeling to establish the RfC, this reviewer believes an UF of 10 is more appropriate than a UF of 3 for the quality of the database.

Mohammad I. Sabri

To account for database deficiencies, an uncertainty factor of 3 is not enough in the derivation of the RfC. The database includes one developmental study with 2-hexanone, where LOAEL but not NOAEL was identified. Although, limited information is available on reproductive, developmental toxicity, neurotoxicity, immunotoxicity of 2-hexanone, an UF of 3 has been applied based on the studies with *n*-hexane and 2,5-HD. To this

reviewer an UF of 10 would be required to account for the database deficiency in the document.

Alan H. Stern

With respect to the UF for database insufficiency, it is stated that the NOAEL of 200 for developmental effects reported for n-hexane in the NTP developmental study is twice the NOAEL identified from the inhalation studies for 2-hexanone in the document. However, the designation of 100 ppm as a NOAEL is based on considering only 6-months of the 10-month Johnson et al. (1977) study (as I have discussed previously). I would argue that 100 ppm in that study is, in fact, a LOAEL given the observations at 9 and 10 months in monkeys (and 25 and 29 weeks in rats). Even leaving this comparison aside, one of the other reviewers (Dr. Abou-Donia?) has noted that although n-hexane is, indeed, metabolized to 2-hexanone, the rate of this metabolism is slow. Therefore, quantitative information on the dose-response relationship for the reproductive and developmental effects of n-hexane are not necessarily useful for assessing the dose-response relationship for reproductive and developmental effects of 2-hexanone – especially given the relatively short time frame of reproductive and/or developmental exposures. Furthermore (and as above for the consideration of developmental effects for the assignment of UFs for the RfD derivation), "aggregated and fused axons" is described as a minimal effect. It is not immediately obvious from this description that this is, indeed, a minimal effect. Information provided by two of the reviewers, Drs. Abou-Donia and Sabri, indicates that fused axons occurs on a continuum of effects that also include axonal swelling and axonal atrophy and probably reduced nerve conduction velocity. In addition, the immunotoxicity considerations raised in the derivation of the RfD are not considered here even though they are derived from inhalation exposures. Each of these reasons combines to suggest that a UF of 10 is most appropriate for addressing database uncertainty.

Bernard Weiss

First, I think the authors should have at least commented on the atmospheric levels of 2-hexanone in the Billmaier et al (1974) study of workers. Neuropathies were detected in this group whose exposure was presumably far below the TLV and approaching the RfC. The document offers some possible explanations, but these belong as well in 5.2.4. They need explanation in context.

Second, the developmental neurotoxicity study cited (Peters et al, 1981) suggests the possibility of a fetal solvent syndrome from 2-hexanone exposure. Unfortunately, the data from the lowest concentration were lost, so only the two highest values contributed data. Enough was seen, however, to warrant some concern about fetal neurobehavioral development. I am not persuaded by the negative effects from Bus et al (1979) and others given their crude measures of function. These represent old-fashioned toxicology.

6. Please provide any other comments on the derivation of the RfC.

Bahie Abou-Donia

The total uncertainty factors should be 3,000 instead of 1,000 as follows:
DU_H = 10
UF_S = 10
UF_A = 10
UF_D = 3

A. John Bailer

Some of the Appendix B-2 summaries are irrelevant. Table B-2.1 list models that were "fit" but no p-values listed. Why were these models included? The only model reported was a linear fit. What was the justification for the particular benchmark response selected for this endpoint?

Frederick J. Miller

In derivation of both the oral and inhalation reference concentrations, the assumption that the product of concentration (C) and time (T) is constant is made referencing U.S. EPA (2002) when discussing adjusting the BMDL to reflect continuous exposure when the principal study actually used less than continuous exposures. In U.S. EPA (2002), a broader model $C^n \times T = k$ is discussed based upon ten Berg (1988). This model is a simplification of the power law family of curves ($C^\alpha \times T^\beta = k$) that Miller et al. (Toxicol. 149:21-34, 2000) showed holds for the family of curves relating concentration and duration of exposure to a fixed level of response for a given endpoint. Defaulting to a value of 1 for n or α and also for β basically invokes Haber's Rule, which Miller et al. (2000) showed does not hold for most biological endpoints. For example, using the large data set for the mouse infectivity model, Miller and colleagues found $\alpha = 0.947$ and $\beta = 0.294$, which shows that C is much more important than T.

Looking at the pulmonary absorption data for 2-hexanone exposure in humans demonstrates that $C \times T = k$ does not hold. Exposure to 100 ppm for 4 hours yields a product of 400 ppm hrs and leads to a breath concentration of 22 ppm of 2-hexanone. Exposure to 50 ppm for 7.5 hours gives a product of 375 ppm hrs, which is close to 400, but results in a breath concentration of only 9.3 ppm.

Whenever possible, the experimental data should be used to determine the exponent on C rather than forcing it to a value of one. In U.S. EPA (2000), an incorrect argument is made that letting n = 1 is being protective of human health. For longer exposure durations like those considered for the RfC, the use of one as the exponent of C over estimates the actual value of C for a given level of effect whenever the true value of the exponent is < 1, which is most often the case. That means that the true value of C is actually lower and that the animal or human would be expected to have the specified effect at a lower level

than what the RfC implies. Moreover, this is the situation for all values of T beyond where the true curve crosses the line of identity.

The Agency is about ready to release for peer review a document covering the model C^n x T = k. Since such models are easy to fit computationally, this reviewer sees no reason that why the actual 2-hexanone data should not be used to determine the exponent of C, if there are sufficient data to do so.

Mohammad I. Sabri

The RfC is an estimate of a continuous inhalation exposure to 2-hexanone that is unlikely to have toxic effects over a lifetime. Neurological effects are a characteristic and sensitive endpoints of inhalation exposure to 2-hexanone, but co-exposure with methyl ethyl ketone (MEK) can significantly potentiate the neurotoxicity of 2-hexanone and thus may influence the RfC.

Alan H. Stern

I have recalculated the BMCLs for the two monkey and two rat endpoints at 6 months and 25 weeks respectively using BMDS version 1.4.1c as opposed to version 1.3.2 used in the document, but employing the same benchmark dose parameters.

Monkey sciatic-tibial

model	**p-value v 1.3.2**	**AIC v 1.3.2**	**BMDL v 1.3.2**	**p-value v 1.4.1c**	**AIC v 1.4.1c**	**BMDL v 1.4.1c**
1st degree polynomial (linear)	0.59	105.59	243.262	0.59	109.590	247.635
2nd degree polynomial		107.29	68.8063		111.296	70.216
power		105.58	243.262	0.59	109.590	247.635
Hill						

Monkey ulnar

model	**p-value v 1.3.2**	**AIC v 1.3.2**	**BMDL v 1.3.2**	**p-value v 1.4.1c**	**AIC v 1.4.1c**	**BMDL v 1.4.1c**
1st degree polynomial (linear)	0.90	105.31	278.471	0.90	109.312	283.714
2nd degree polynomial		107.29	94.1174		111.296	96.039
power		109.31	278.471	0.90	109.312	283.714
Hill						

Rat sciatic-tibial

model	**p-value v 1.3.2**	**AIC v 1.3.2**	**BMDL v 1.3.2**	**p-value v 1.4.1c**	**AIC v 1.4.1c**	**BMDL v 1.4.1c**
1st degree polynomial (linear)	0.79	121.55	232.105	0.79	123.557	237.78
2nd degree polynomial		123.48	100.045		125.483	103.14
power		125.48	232.56		125.483	248.26
Hill						

Rat ulnar

model	**p-value v 1.3.2**	**AIC v 1.3.2**	**BMDL v 1.3.2**	**p-value v 1.4.1c**	**AIC v 1.4.1c**	**BMDL v 1.4.1c**
1st degree polynomial (linear)	0.26	122.77	352.274	0..26	124.776	360.84
2nd degree polynomial		606.18	87.3758		125.483	64.539
power		125.48	14.1388		124.777	360.844
Hill						

For both monkey models and for the rat sciatic-tibial models the values obtained with the newer version of the BMDS software are similar to, but the BMDL values are larger than those obtained with the older version of the software. For the rat ulnar models, however, there are some significant differences in the BMDLs as well as in some of the AIC values. It is not clear to me what the source of these differences is.

In addition, as I have previously stated in response to General Charge Question 4, I believe that the question of whether benchmark dose modeling is appropriate for these data at all is important and has not been dealt with in the document. This question goes not only to whether there are sufficient data to support a benchmark dose approach, but also whether the necessity of the of having at least two dose levels in addition to the control precludes using the later stage data for which the LOAEL is 100 ppm rather than the value of 1,000 ppm in the time point used for the benchmark dose modeling.

Bernard Weiss

I think they were well-reasoned.

(D) Carcinogenicity of 2-hexanone

1. Under the EPA's 2005 Guidelines for carcinogen risk assessment (www.epa.gov/iris/backgr-d.htm), the Agency concluded that there is inadequate information to assess the carcinogenic potential of 2-hexanone. Please comment on the scientific justification for the cancer weight of the evidence characterization.

Bahie Abou-Donia

The EPA's conclusion is correct and justified.

A. John Bailer

Since data on carcinogenicity are not available in either humans or animals, "inadequate information" is the only conclusion possible.

Frederick J. Miller

The Agency is completely justified in concluding that there is inadequate information to asses the carcinogenic potential of 2-hexanone.

Mohammad I. Sabri

There are no published studies that adequately provide the information on carcinogenic or genotoxic potential of 2-hexanone. Occupational studies also do not provide any evidence for carcinogenic effects of 2-hexanone. EPA has concluded that there is inadequate information to assess the carcinogenic potential of 2-hexanone.

Alan H. Stern

Given the absence of direct or indirect data on carcinogenicity, this is the only possible categorization.

Bernard Weiss

The current data on hexanone, from my standpoint, do not indicate that it poses a cancer threat. "Inadequate" is perhaps too neutral, although I appreciate the constraints resulting from EPA terminology.

V. SPECIFIC COMMENTS

Bahie Abou-Donia

2-Hexanone is a solvent that has been shown, both in humans and several animal species, to produce central-peripheral distal axonopathy, characterized by the presence of giant axonal swilling and degeneration followed by demyelination, as the main toxic action. In addition, 2-hexanone is a potent inducer of many cytochrome P-450 isozymes that led to its ability to enhance the neurotoxic action of other chemicals such as organophosphates. Furthermore, 2-hexanone-induced cytochrome P-450 would result in increasing the metabolism of other chemicals such as therapeutics, leading to their diminished efficacy. This may require an additional uncertainty factor as a modified factor.

A. John Bailer

[p. 1] typo? "#24 hours" vs. "<24 hours"?

[p. 22] LD50 reported but "details for either study are limited." If these details are limited, then how much attention should be devoted to these values? What if these were based on a small number of animals?

[p. 23] Insufficient to report average response levels without some indication of variability and number of animals on test, i.e. additional information should be added to Table 4-3.

[p. 28] Table 4-5 gives pup-specific summaries although litter-specific summaries/analyses are usually recommended because of the lack of independence of pups in a litter

[p. 30] So, a 0-1-2 score was assigned to each of 10 criteria resulting in a 0 to 20 cumulative score that was analyzed using a rank-based procedure, Kruskal-Wallis test. The mean (sd) ranks are presented in Table 4-6. Is this a meaningful display? Wouldn't the summary statistics from the 0-20 scale score be more informative? By the way, was the "p-value" summarizing an analysis within each time period or across all time periods? This is unclear.

[p. 34] Table 4-8 is clearly one of the most important tables in the report. It demonstrates strong dose-response relationship that will be used in later risk-based estimates. The table legend could be clarified. (This is true for most of tables.) The table legend should note that cell entries are the number of animals exhibiting the response/number of animals exposed to treatment.

[p. 50] Interesting question to think about 2-hexanone as a chemical that can "potentiate the neurotoxic effects of EPN." What if that was the only effect of this chemical? How would this be accommodated in the IRIS database?

[p. 51] Table 4-15 – not all NOAELs/LOAELs are created equally! n=3 vs. n=10 along with intrinsic variability in test systems may dramatically impact the power to detect a difference between experimental groups. Why not report BMD/BMDL for studies if the data are available?

[p. 53] Table 4-16 – what does "adjusted" mean as used in the last row of this table?

[p. 60] Table 5-2 – report more details about the model fit? Does the quantal quadratic take the form Prob = bkgd + (1-bkgd)*{1-exp(– q2*d^2)}? What multistage model was fit? Prob = bkgd + (1-bkgd)*{1-exp– q1*d -q2*d^2-q3^d^3)}? How was the "p-value" constructed? Pearson X^2 goodness-of-fit statistic? Models reported (along with the models reported in the Appendix) all provided a good fit to the data. No real surprise here since lots of parameters relative to the number of different doses tested (e.g. 3 parameters in multistage model for a 4 treatment study).

[p. 60] Comment that the assumption that the product of dose-time constant scaling is also commonly known as Haber's Rule.

[p. 67] The strategy of working from the F statistic back to the pooled variance estimate was reasonable ASSUMING that the variance is common in the different treatment group. This is a strong assumption for many studies where variability tends to increase with mean responses. Are you confident that equal variability in equal treatment group is a reasonable assumption for this response?

[p. 68] Table 5-3 – report sample sizes in table heading (n_i=8?). Include the estimated common standard deviation? I would like to see a figure as well.

[p. 69] You are selecting a model using an AIC criterion. More description on the models fit to the continuous responses should be included here.

[p. 70] Do partition coefficients have a distribution across a population?

[p. 75] Display with log10 y-axis make the RFCs appear less variable than they are (more than a 10 fold difference).

[p. 78] Include at least a partial summary of the TOXNET values. How much do these differ from the proposed RfD/RfC?

[p. 80] Report magnitude of decrease used to define the RfC.

[Appendix B-1] Almost impossible to read. Should be produced using a FIXED width font if you are going to report computer output. Note that you are fitting a 4 parameter model to 4 data points [B-3]. You set parameter convergence criteria to 10^(-8) and you estimate of beta3 was 3.7x10^(-8). Do you have any concerns about convergence issues? Should you redo this with tighter convergence criteria? You could also rescale the doses and this would no longer be an issue, e.g. d_new = d/(max d). Was the AIC

adjusted for the number of parameters fit or the number of parameters not on the boundary?

[Appendix B-2] Why are missing values in Table B-2.1? Clarify the model being fit – e.g. what is "alpha"?

[B-10] A LR stat = 0 with a P-value < .0001? This doesn't make any sense.

[B-11] The figure is almost impossible to read.

Frederick J. Miller

- The treatment of tabular data is inconsistent; error terms are provided in many tables but whether they are standard errors or standard deviations is not always made clear. In addition, table legends should specify the sample sizes upon which the means are based whenever possible. Basically, all tables should be able to stand alone.

- In Table 4-11, p-values of 0.01 and 0.05 are noted. This is really overkill because there is statistical literature showing how large the computed F-statistic needs to be before one can really claim that something is really significant at a lower probability level than what the investigators were using.

- On page 49, can the authors indicate if the co-exposure studies discussed there are likely to represent possible co-exposure situations for humans? If so, extra caution is needed.

- On page 69, the 2nd paragraph should refer to Table 5-5 and not Table 5-4. Also, the reference should be to Appendix B-2 and not to appendix B-1.

- On page 78, the paragraph referring the reader to a web site that has assessments by other organizations should be expanded. This reviewer went to the web site and found it difficult to retrieve any useful information.

Mohammad I. Sabri

The draft document has reviewed the available information on the toxicity/neurotoxicity of 2-hexanone. Although, the overall confidence in the two principal studies (O'Donoghue et al., 1978 and Johnson et al., 1977) is medium, the review has satisfactorily described and discussed relevant literature to derive the RfD and RfC of 2-hexanone. Because of inadequate information, the carcinogenic potential of 2-hexanone is not unquantifiable.

Corrections:

1. Page 58, 2nd para, reference Abdo et al. (1982) should be Abou-Donia et al. (1982)?

2. Page 62, section 5.1,5 line 13, the reference Abdo et al., 1982 should be Abou-Donia et al., 1982?

3. Page 76, line 4, section 4.6.1 should be section 4.7.1?

Alan H. Stern

Pg. 6, par. 2 - "*The quantities of 2-hexanone absorbed by tow volunteers were …"* Does this refer to systemic absorption, or simply to retention in the stratum corneum.

Pg. 9, par. 3 - *"However, with both routes of administration, 2-hexanol concentration did not appear to be dose dependent."* This seems strange, and warrants further discussion.

Pg. 15, par. 3 - *"Because inhalation exposure of humans to 1-[^{14}C]-2-hexanone resulted in the appearance of labeled carbon dioxide…hypothesized that the metabolic pathway…is similar in humans and experimental animals."* Why? This appears to be a non-sequitur.

Pg. 16, par. 4 - 20 μg is referred to here as a concentration. This should be 20 μg/L.

Pg. 24, par. 1 - *"Other than neural effects and changes in body weight, no non-neural clinical signs… were found."* What about changes in liver, kidney and testes relative weight?

> It can be pointed out that the lack of effects seen with a dose of 50 mg/kg/day of MiBK was 13 times the concentration present as a contaminant in the two previously referenced studies.
>
> *"O'Donoghue et al. (1978) did not observe adverse effects in the kidney or liver…"* Again the increased relative weight is not being addressed. If, indeed, this is not considered an adverse effect, this should be explicitly stated and the rationale for such a conclusion should be discussed.

Pg. 31 - If no locomotor activity for the 0.1% exposure was reported by the authors in Abdel-Rahman et al. (1978), this should be specifically stated.

Pg. 34, par. 3 - The reference in the last sentence on the page to an exposure of 60 ppm for 6 months does not appear to be linked to any study or exposure previously discussed.

Pg. 38, par. 3 - I do not believe that the acronym "MCV" has been previously defined.

Pg. 40, par. 1 - Although there is a low dose (100 ppm) in this study, only the high dose and control results are discussed.

Pg. 43, par. 2 - "*In order to explore the effects of 2-hexanone…"* The interpretation and hypothesis testing here are unclear.

Pg. 50, par. 3 - I don't see the Krasavage et al. (1980) data discussed elsewhere in the document. It is important to discuss this in detail since comparison of the proposed RfD to the RfD for n-hexane would be very useful.

Pg. 52, par. 1 - Should note the contribution from dermal exposure in this cohort.

Fig. 4-1 - It would be helpful to show the ring-closure step.

Pg. 57, par. 2 – *"…during development may likely"* This language does not make logical sense.

Pg. 58, par. 2 - The text of the Abdel-Rahman et al. (1978) study does not report data for the 0.1% exposure (the low dose). This should be reflected in the rationale in the document for not using this study as the principal study.

Pg. 60, par 2 - The formula and calculations for the derivation of the $BMDL_{ADI}$ are unnecessary given that, in fact, no adjustment is made.

Pg. 65, par. 3 - It should be noted that for Egan, Krasavage and O'Donoghue (1977 and 1979) and Johnson, the lowest concentration was the same, 100 ppm. For Katz, the lowest concentration was 70 ppm

Pg. 66, par. 2 - More information on the specification of the BMD model would be useful here (e.g., how was the BMR specified? What was the rationale?).

Pg. 68, par. 1 - The estimates produced by this procedure should be presented here. Also, the n's for each dose group and time point should be given here.

Pg. 69 - There is no explanation of how the various time points were used in the benchmark modeling. While I was able to eventually deduce this, this should be stated explicitly

> Par. 1 - A first degree polynomial is a linear model. Why not just use the term, "linear?"

Pg. 70, par. 2 - Why weren't data modeled as mg/m^3 to begin with?

> Par. 4 - The classification of 2-hexanone as a category 3 gas needs to be justified in terms of the definition of a category e chemical.

Bernard Weiss

I want to reinforce my previous recommendation that each section begin, in essence, with an introductory paragraph that guides the reader to the section's conclusions.