# Peer Review Workshop for EPA's Draft Toxicological Review of Thallium and Compounds

# Reviewer Post-Meeting Comments

Submitted to:

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC 20460

Submitted by:

Eastern Research Group, Inc.
110 Hartwell Avenue
Lexington, MA 02421

May 29, 2008

# Contents



## Reviewer Post-Meeting Comments



## Appendices

# Introduction

The Integrated Risk Information System (IRIS) is a U.S. Environmental Protection Agency (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: hazard identification and dose-response evaluation. IRIS information includes a reference dose (RfD) for noncancer health effects resulting from oral exposure, a reference concentration (RfC) for noncancer 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.

EPA's Office of Water, in collaboration with EPA's National Center for Environmental Assessment (NCEA), developed a *Draft Toxicological Review of Thallium and Compounds*. This assessment updates assessments for thallium compounds currently available on the IRIS database. Thallium was nominated for IRIS reassessment by the Office of Water in light of new data available on the reproductive/developmental toxicity of thallium potentially relevant to the compound's National Primary Drinking Water Regulation (NPDWR) six-year review; this chemical was added to the IRIS agenda for assessment development in 2002.

In May 2008, Eastern Research Group, Inc. (ERG), an EPA contractor, organized an independent peer review of EPA's *Draft Toxicological Review of Thallium and Compounds*. The review document contained a chronic RfD for soluble thallium salts and a qualitative cancer assessment. Assessments were also provided for two thallium compounds for which the available literature was insufficient to develop reference values. ERG identified six nationally recognized experts (Appendix A) to conduct this review:

- Ronald Baynes, North Carolina State University
- George Cherian, University of Western Ontario
- Lucio G. Costa (chair), University of Washington
- George Daston, The Procter & Gamble Company
- Robert Hoffman, New York University
- Deborah Rice, Maine Center for Disease Control

ERG provided the reviewers with a charge (page 3), which asked for their comments on the various aspects of the document. Reviewers were also provided with the study used as the principal basis of the RfD (MRI [Midwest Research Institute], 1988)[1] and with a compilation of

[1] MRI (Midwest Research Institute). 1988. Toxicity of thallium(I) sulfate (CAS No. 7446-18-6) in Sprague-Dawley rats. Volume two: Subchronic (90-day) study. Revised final report. Project No. 8702-L(18), Work Assignment No. 111148-008. Prepared for U.S. Environmental Protection Agency, Office of Solid Waste, Washington, DC, through

comments on this study from a 2006 peer review (ORNL, 2006)[2]. Each reviewer also received complete copies of the written comments submitted during the public comment period, which they were asked to consider.

In the first stage of the review, the experts worked individually to prepare written pre-meeting comments, which were provided to all reviewers and EPA prior to a one-day peer review workshop. In the second stage, ERG convened the one-day workshop, on May 19, 2008, at a venue in Arlington, Virginia. The meeting was open to the public and attended by 3 observers and an ERG facilitator (Appendix B). Appendix C provides the workshop agenda. The meeting format included an opportunity for public comment. No members of the public were present to comment. After the meeting, reviewers revised their pre-meeting comments to reflect their views as they had evolved based on the workshop discussions. The reviewer final post-meeting comments are provided in this report. These comments reflect the individual opinions of the reviewers.

Dynamac Corporation, Rockville, MD.

[2] ORNL (Oak Ridge National Laboratory) 2006. External Letter Peer Review of a Report by Midwest Research Institute, Revised Final Report: Toxicity of Thallium (I) Sulfate (CAS No. 7446-18-6) in Sprague-Dawley Rats, Volume Two: Subchronic (90-day) Study, July 1988, Compilation of Reviewer Comments and Responses to Charge Questions. Prepared for the EPA by ORNL under interagency agreement No. DW-89922097 between EPA and the U.S. Department of Energy. DOE Project No. 1824-S881-A1.

# Peer Review Workshop for Toxicological Review of Thallium and Compounds

U.S. Environmental Protection Agency
Task Order No. 22
Contract No. EP-C-07-024

## TECHNICAL CHARGE TO PEER REVIEW PANEL

### Background

The U.S. Environmental Protection Agency (EPA) is seeking an external peer review of the scientific basis supporting the human health assessment of thallium and compounds that will appear on the Agency=s online database, the Integrated Risk Information System (IRIS). Eastern Research Group (ERG), under contract to EPA, is managing this external peer review. IRIS is prepared and maintained by the EPA's National Center for Environmental Assessment (NCEA) within the Office of Research and Development (ORD). Existing IRIS assessments of selected thallium compounds were posted to the database in 1987.

The draft health assessment includes a Reference Dose (RfD) and a carcinogenicity assessment. Below are a set of charge questions that address scientific issues in the assessment of thallium and compounds. Please provide detailed explanations for responses to the charge questions.

As the study used for the principal basis of the RfD (Midwest Research Institute [MRI], 1988) was not peer reviewed, EPA had an external review of the study conducted in 2006. To help inform your evaluation, EPA is also providing you with the external peer review report summary. The MRI study can be accessed from EPA's website at: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=188304

### Charge Questions

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 thallium and thallium compounds.

3. Please discuss research that you think would be likely to increase confidence in the database for future assessments of thallium and compounds.

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:

(A) Oral reference dose (RfD) for thallium

A1. The 90-day oral gavage study by Midwest Research Institute (MRI, 1988) was selected as the basis for the RfD. 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.

A2. Alopecia (hair loss) was selected as the most appropriate critical effect for the RfD. EPA characterizes alopecia as being an adverse effect. Please comment on whether the science and mode of action information supports alopecia as an adverse effect. EPA has stated: "Whether alopecia is itself an adverse effect merits consideration. In humans, alopecia is generally reversible upon cessation of thallium exposure. Alopecia, however, appears to be a part of a continuum of dermal changes observed following thallium exposure, as well as one of a spectrum of effects on target organs that include the nervous and gastrointestinal systems. For these reasons, alopecia supported by two cases of hair follicle atrophy is considered an adverse effect." Please comment on whether the selection of this critical effect has been scientifically justified. Is EPA's choice transparently and objectively described in the document? Please provide a detailed explanation. Please identify and provide the rationale for any other endpoints that should be used instead of alopecia to develop the RfD.

A3. At the high dose in the MRI (1988) study, two female rats exhibited moderate to severe alopecia that could not be attributed to self-barbering or normal cyclic hair growth. Histologic examination of skin samples from these high-dose females showed atrophy of hair follicles. EPA considered these findings to be adverse, and thus the high dose in this study (0.25 mg/kg-day thallium sulfate) to be the lowest-observed-adverse-effect level (LOAEL). The mid-dose group (0.05 mg/kg-day thallium sulfate) was identified as the no-observed-adverse-effect level (NOAEL). Is EPA's interpretation of the study findings scientifically justified? Has this interpretation of the findings been transparently and objectively described in the document?

As part of the evaluation of alopecia as a critical effect for the RfD, EPA performed a series of Fisher's Exact Tests to determine if the incidence of alopecia in any of the three dose groups was statistically significantly elevated above controls using all cases of alopecia reported by MRI (1988). Please comment on whether EPA chose the appropriate data set and the appropriate statistical test for this analysis.

The study investigators reached a different interpretation of the study findings than did EPA. The investigators considered alopecia to be attributable to the cyclic pattern of hair growth in rodents and, consequently, did not consider these findings to be biologically significant. The high dose (0.25 mg/kg-day thallium sulfate) was identified in MRI (1988) as the NOAEL. Is the study authors' conclusion that the high dose (0.25 mg/kg-day thallium sulfate) represents a NOAEL justified and supported by the study data?

A4. The traditional NOAEL-LOAEL approach was used to define the point of departure (POD) for the RfD. A benchmark dose (BMD) analysis was considered but was not conducted because of the nature of the data set for alopecia. Please provide comments with regards to whether a NOAEL-LOAEL approach is the best approach for determining the POD. Has the approach been scientifically justified? Is it transparently and objectively described? Please identify and provide a rationale for

any alternative approaches for the determination of the POD, and if such approaches are preferred to EPA's approach.

A5. Please comment on the selection of the uncertainty factors applied to the POD for the derivation of the RfD. For instance, are they scientifically justified and transparently and objectively described in the document? If changes to the selected uncertainty factors are proposed, please identify and provide a rationale(s).

A6. Please comment specifically on the database uncertainty factor of 10 applied in the RfD derivation. Please comment on the use of the database uncertainty factor specifically for the lack of adequate developmental toxicity studies and a two-generation reproductive toxicity study, and additional uncertainty associated with the limited data available on neurotoxicity in light of the potential for neurotoxicity to represent a sensitive endpoint for thallium exposure. 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?

(B) Inhalation reference concentration (RfC) for thallium and compounds

No data are available to derive the RfC for thallium and compounds. The only published studies involving inhalation exposure include a few case reports (Hirata et al., 1998; Ludolph et al., 1986) that suggest an association between occupational exposure and toxicity (including alopecia, gastrointestinal symptoms, and neuropathy), but the route or routes of exposure in these workplace setting could not be established.

B1. Has the rationale and justification for not deriving an RfC for thallium been transparently described in the document?

(C) Carcinogenicity of thallium and compounds

C1. 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 thallium and compounds. Please comment on the scientific justification for the cancer weight of the evidence characterization. A quantitative cancer assessment was not derived for thallium. Has the scientific justification for not deriving a quantitative cancer assessment been transparently and objectively described?

**Ronald Baynes, DVM, Ph.D.**
North Carolina State University
Raleigh, NC 27606

Dr. Baynes received a B.Sc. in 1984 from the University of the West Indies, Barbados, a DVM in 1990 from Tuskegee University, an MS in Pharmacology from the University of Georgia in 1992, and a PhD in Pharmacology from North Carolina State University in 1997. He is currently an Associate Professor of Pharmacology at North Carolina State University in the Center for Chemical Toxicology Research and Pharmacokinetics (CCTRP) at the College of Veterinary Medicine and Department of Toxicology in the College of Agricultural and Life Sciences. Dr. Baynes has over 15 years experience in pharmacokinetics and chemical risk assessment of chemical mixtures as it pertains to dermal absorption. Dr. Baynes has prepared Toxicological profiles for ATSDR and summaries for US EPA. He has been a PI for the last 8 years on a NIOSH funded research project focused on using in vitro skin and quantitative models to assess the dermal absorption of chemical mixtures that are of occupational concern. Additionally, he is Co-Director of the USDA-funded Food Animal Residue Avoidance Databank (FARAD) that develops classical and novel models (e.g., PBPK, and popPK models) to assess the depletion of drug and chemical residues in food animals to minimize the human exposure to these chemicals in their diet. Dr. Baynes also manages the chemical analytical lab in the CCTRP at North Carolina State University and supervises several staff and graduate students under his mentorship. Dr. Baynes has also written more than 60 peer-reviewed publications and book chapters on dermal absorption kinetics and pharmacokinetics in general. He is a member of Sigma Xi, American Chemical Society, Society of Toxicology, and the American Veterinary Medical Association to mention a few.

**General Charge Questions:**

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

This US EPA document attempts to present the scientific evidence for noncancer and cancer hazard of thallium. The document also describes the conclusions by MRI (1988) which attributed alopecia to the cyclic pattern of hair growth. The US EPA case for thallium-induced alopecia could have been strengthened if they discussed the hair growth cycle in rats and how a clinical condition such as alopecia can be distinguishable from hair loss associated with normal growth cycle in rats. In general, the document is weak in accurately providing a quantitative assessment based on other adverse endpoints identified in the MRI (1988) study. To basically ignore these other endpoints and focus on alopecia questions the objectivity of the assessment.

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

There is a paucity of other suitable studies that should be considered for this assessment. There are other studies (e.g., Wei et al., 1987) only briefly mentioned in this EPA assessment that could have been used to calculate a NOAEL or BMDL. Based on the current EPA assessment, there is a need to better characterize the dose-response relationship between thallium and alopecia in future studies. This needs to be accompanied with histopathological examination of affected skin. There are knowledge gaps and inconsistencies regarding how soon after exposure that laboratory animals can be diagnosed as expressing alopecia.

3. ***Please discuss research that you think would be likely to increase confidence*** **in the database for future assessments of thallium and compounds.**

The following research activities would greatly improve our understanding of thallium toxicity and would provide at the very least dose-response relationships that can utilized in a more quantitative and mechanistic based risk assessment with significantly reduced uncertainty:

i. Establish a more accurately defined dose-response relationship with appropriate dose ranges that capture the alopecia endpoint. This would include supportive histopathological examination of alopecia skin and apparently unaffected skin areas of all treated and untreated animals and supportive in vivo and in vitro experiments demonstrating thallium interaction with hair follicle at various stages of the hair cycle in both males and females.
ii. With the appropriate doses, conduct a chronic 2-year study to identify neurological, reproductive, endocrine, and cardiovascular endpoints. For example, how does thallium affect, if at all, the pituitary-hypothalamus axis? The alopecia may be a biomarker of endocrine disruption.
iii. Conduct acute dose-range finding studies and if needed chronic inhalation study to identify the hazard of inhaling thallium.

iv. Establish whether absorption, distribution, and elimination of thallium is linear following oral and dermal absorption of thallium salts. There is a very casual assumption in the open literature that oral and dermal bioavailability is almost 100%. Physicochemical description of this chemical suggests otherwise.
v. Using the above databases to conduct appropriate PBPK modeling to better understand thallium dosimetry and obtain a BMDL and RfD with significantly reduced uncertainty.

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?**

For the most part, several of the UFs have been adequately described in this review by US EPA. This reviewer only questions the UF for subchronic to chronic extrapolation as described below. This document should really provide sound reasoning for why the UF has been reduced from 10 to 3 for subchronic to chronic extrapolation. As the document reads on page 63, alopecia is recognized as occurring within weeks so no need for subchronic to chronic extrapolation. This begs the question then that the UF for this extrapolation should be ONE (1) instead of 3. EPA needs to expand on this uncertainty.

**Chemical-Specific Charge Questions**:

**(A) Oral reference dose (RfD) for thallium**
***A1. The 90-day oral gavage study by Midwest Research Institute (MRI, 1988) was selected as the basis for the RfD. Please comment on whether the selection of this study as the principal study has been scientifically justified.***

Based on all available data, the MRI (1988) study appears to be the most appropriate study to base the derivation of an RfD. *In vivo* toxicology data for any form of thallium is limited, but better described for Thallium I salts such as sulfate and nitrates. The former was described in the MRI (1988) study for oral gavage in a subchronic study. This is supported by anecdotal reports and several incomplete epidemiological data sets where thallium levels in the hair and urine (<0.1 to 76.5 ppb) were higher than back ground or reference population levels (0.1 to 1.2 ppb). A major weakness of this epidemiological study by Brockhaus et al., (1981) was the negative correlation between % alopecia and thallium conc in the urine and hair. With incomplete histories of time of exposures and exposure levels, this observation is not surprising. The major weakness with the MRI (1988) study is that histopathological examinations of skin were only completed for the vehicle control and the highest dose (**0.2 mg/kg dose**) and not the **control, 0.008 or 0.04** mg/kg dose groups. This is most critical as the NOAEL was based on the 0.04 mg/kg dose where there was a ***statistically significant increase in incidence of alopecia*** but skin from this dose group was not examined for hair follicle atrophy. Alopecia with hair follicle atrophy was observed with the high dose of 0.2 mg/kg which strongly suggest thallium interaction with hair growth;

but this was assigned the LOAEL. The EPA document needs to define alopecia with atrophy as the adverse effect and not just alopecia. Note that the incidence of alopecia in both male and female rats at 0.04 mg/kg was determined to be statistically significant, yet this dose of 0.04 mg/kg was deemed to be the NOAEL. Why isn't this the LOAEL?

This reviewer believes that alopecia is an adverse event and this is supported by thallium induced telegen (resting) follicles in dogs with thallium poisoning and the MRI study. There are other in vivo studies that support these findings at even higher doses than the MRI study. For example one study identified the NOAEL and LOAEL for alopecia as **0.4 and 1.2 mg/kg** BW, respectively, thallium acetate (Downes et al., 1960). However, mortality was too high in control and treatment groups for this study to be considered as the principal study for derivation of the RfD. Another study (Manzo et al., 1983) demonstrated alopecia and nerve histology and more importantly, 15-20% mortality at **1.4 mg/kg** for 36 weeks which is 7 times the MRI (1988) study. This study tested only one dose, so it is likely that doses significantly less than the **1.4 mg/kg** could have identified a NOAEL not only for alopecia, but also for nerve histology. El-Garaway et al. (1990) administered only one dose of **0.65 mg/kg for 90 days** and observed an increasing BUN, Cr, ALT, and Bilirubin. While all of these studies do not provide a dose-response relationship, but taken together there is some credible evidence that the 0.2 mg/kg LOAEL from the MRI (1988) study may not be a overly conservative estimate of the LOAEL for a critical endpoint other than alopecia.

***Has this study been transparently and objectively described in the document?***

For the most part, the MRI (1988) study has been adequately described in the EPA document. The authors could have provided a better time line for the appearance of alopecia as well as hair loss associated with the normal hair cycle in the rat. The distribution pattern of the alopecia could have been better described as well in order to assess reproducible patterns that can be attributed to thallium. This reviewer had to glean this from the original MRI (1988) study which should be the case in a risk assessment document such as this. There are other toxicological endpoints (e.g., LDH, exophthalmos, etc) in this MRI study that were barely mentioned but no NOAEL or BMDL was attempted even though statistically significant changes were reported in Table 3 page 19.

The exophthalmos and lacrimation endpoints need to be better explained and why they were not considered in this risk assessment. Ophthalmic and histopath exam of the eye revealed no remarkable effects. There is no mention of whether a statistical analysis was conducted with this data and endpoints that appear to be more adverse than alopecia. However, both the alopecia and exophthalmos are

associated with hyperthyroidism. No evidence that TSH or thyroxine levels were assessed in any of the rodent in vivo or human epidemiological studies.

***Please identify and provide the rationale for any other studies that should be selected as the principal study***.

There are several other *in vivo* studies that could have been selected as the principal study for RfD derivation. However, EPA provided limited justification why these studies were not included. EPA only explained that the Downes et al. (1960) study had significant mortality and the Wei (1987) reproductive study did not follow EPA protocols. In the case of the latter study, this reasoning is not sufficient to explain its exclusion and for not calculating a NOAEL or BMDL. Epidemiological studies would be ideal, but all available epidemiological studies provided limited evidence of cause and effect associated with exposure in food or inhalation.

**A2. Alopecia (hair loss) was selected as the most appropriate critical effect for the RfD. EPA characterizes alopecia as being an adverse effect. *Please comment on whether the science and mode of action information supports alopecia as an adverse effect.* EPA has stated:"Whether alopecia is itself an adverse effect merits consideration. In humans, alopecia is generally reversible upon cessation of thallium exposure. Alopecia, however, appears to be a part of a continuum of dermal changes observed following thallium exposure, as well as one of a spectrum of effects on target organs that include the nervous and gastrointestinal systems. For these reasons, alopecia supported by two cases of hair follicle atrophy is considered an adverse effect." Please comment on whether the selection of this critical effect has been scientifically justified. Is EPA's choice transparently and objectively described in the document? Please provide a detailed explanation.**

Alopecia with histopathology of the hair follicle should be considered an adverse effect and clearly stated as a POD by the EPA document. Again, a significant incidence of alopecia was observed at the assigned NOAEL dose of 0.04 mg/kg. Unfortunately histopathology was only observed in the assigned LOAEL dose and not examined in the assigned NOAEL and lower doses and there is no evidence that the skin sections examined were from the alopecia areas of the skin. The blood chemistries from the same study demonstrated some minor changes in SGOT and LDH; however, these changes are not organ specific for any toxic event, although suggestive of thallium-induced hepatic effects and a possible dose-response relationship may exist.

**Please identify and provide the rationale for any other endpoints that should be used instead of alopecia to develop the RfD**.

Data from limited several sub-chronic studies (Formigili et al., 1986; Wei, 1987) and in vitro studies (Gregotti et al., 1992) suggest that reproductive (especially male) endpoints are worthy of further investigation. However, these studies were poorly reported by the authors and dose-response relationship is difficult to ascertain. This is however a poor excuse for not attempting to determine a dose-response relationship and providing NOAELs and BMDL. The IP and SQ routes have demonstrated neurotoxicity

at oftentimes high does of thallium (5-50 mg/kg). However, a more recent study by Galvan-Arzate et al. (2000) describes a dose-related increase in thallium deposition in the brain and lipid peroxidation in male Wistar rats after 30 day exposure to 0.8 mg/kg or 1.6 mg/kg dose.

Although EPA reports that there are no studies reporting the effects of chronic oral thallium on the nervous system, there is one study in this review (Manzo et al., 1983) which describes in some detail functional and histopathological changes in PNS and changes in motor and sensory action potentials and axonal destruction by Wallerian degeneration following oral exposure to thallium sulfate in drinking water at 1.4 mg/kg for 36 weeks. No other doses were mentioned in this study which was obtained from a book chapter!! EPA should really provide a more critical review of this paper as neurological effects are often associated with thallium toxicity.

**A3. At the high dose in the MRI (1988) study, two female rats exhibited moderate to severe alopecia that could not be attributed to self-barbering or normal cyclic hair growth. Histologic examination of skin samples from these high-dose females showed atrophy of hair follicles. EPA considered these findings to be adverse, and thus the high dose in this study (0.25 mg/kg-day thallium sulfate) to be the lowest-observed-adverse-effect level (LOAEL). The mid-dose group (0.05 mg/kg-day thallium sulfate) was identified as the no observed-adverse-effect level (NOAEL). Is EPA's interpretation of the study findings scientifically justified? Has this interpretation of the findings been transparently and objectively described in the document?**

The EPA review has provided supporting evidence from acute (poisoning) oral exposures in dogs (Thomas and McKeever, 1993; Schwartzman and Kirschbaum, 1961) that resulted in adverse affects to the skin as well as the hair follicle. The induction of telegen follicles in 13 dogs in the latter study strongly supports the MRI study as the principal study. Further examination of the MRI (1988) range-finding study provided "shedding/balding/erythema" data which could be described as a dose-response from 0 to 5 mg/kg thallium with alopecia appearing within 6-13 days.

However, EPA needs to clearly define what is a NOAEL vs a LOAEL as it pertains to alopecia. Does a dose of 0.04 mg/kg which causes significant incidence of alopecia with no evidence of histopath change in skin/hair follicle a NOAEL as the document suggest? If so, then the document need to make this objectively clear.

**As part of the evaluation of alopecia as a critical effect for the RfD, EPA performed a series of Fisher's Exact Tests to determine if the incidence of alopecia in any of the three dose groups was statistically significantly elevated above controls using all cases of alopecia reported by MRI (1988). Please comment on whether EPA chose the appropriate data set and the appropriate statistical test for this analysis**.

The EPA used a correct nonparametric statistical analysis of the categorical data for which Fisher's Exact Test is well suited. In this case it has determined whether there is some association between the categorical data

**The study investigators reached a different interpretation of the study findings than did EPA. The investigators considered alopecia to be attributable to the cyclic pattern of hair growth in rodents and, consequently, did not consider these findings to be biologically significant. The high dose (0.25 mg/kg-day thallium sulfate) was identified in MRI (1988) as the NOAEL. Is the study authors' conclusion that the high dose (0.25 mg/kg-day thallium sulfate) represents a NOAEL justified and supported by the study data?**

Again, the EPA needs to clearly define what is a NOAEL vs a LOAEL as it pertains to alopecia. Does a dose of 0.04 mg/kg which causes significant incidence of alopecia with no evidence of histopath change in skin/hair follicle a NOAEL as the document suggest? This issue here is more a matter of better identification of the POD as it pertains to alopecia. How does this deviate from its Guidance documents?

EPA has correctly identified an adverse effect that cannot be "attributable to the cyclic patter of hair growth in rodents". The evidence based on the observations and data analysis suggests that alopecia was more likely to be present in thallium exposed rats than rats receiving the vehicle only. While histopathological examination of alopecia areas of the skin from all dose groups would be ideal, there is significant evidence from other acute oral studies to suggest that hair follicle atrophy is strongly associated with thallium exposure based on proposed thallium interaction with organosulfurs in hair as a plausible mechanism for alopecia. Lipid peroxidation in skin and nervous tissue are other proposed mechanisms which should warrant a look at other toxic endpoints in these organs at low doses. While chemically induced alopecia may be reversible, this may not be the case with chemically induced CNS or PNS effects. EPA should also recognize in their review that there are possible estrogen receptor pathways within the dermal papilla that regulates the telogen-anagen follicle transition and that diffusible factors associated with the anagen follicle influence cell proliferation in the epidermis (Ho and Smart, 1996). This may explain male vs. female differences observe din the MRI study. The sex differences were highlighted by one public comment reviewer as a reason to dismiss this alopecia end point as and the EPA-proposed NOAEL as a POD.

**A4. The traditional NOAEL-LOAEL approach was used to define the point of departure (POD) for the RfD. A benchmark dose (BMD) analysis was considered but was not conducted because of the nature of the data set for alopecia. Please provide comments with regards to whether a NOAEL-LOAEL approach is the best approach for determining the POD. Has the approach been scientifically justified? Is it transparently and objectively described? Please identify and provide a rationale for any alternative approaches for the determination of the POD, and if such approaches are preferred to EPA's approach.**

Based on the paucity of dose-response data sets, the NOAEL-LOAEL approach is the most appropriate method to define the POD in this risk assessment as it pertains to alopecia. However, there are other end-points (examples described earlier) in the MRI study that should be considered for BMD modeling approaches.

**A5. Please comment on the selection of the uncertainty factors applied to the POD for the derivation of the RfD. For instance, are they scientifically justified and transparently and objectively described in the document? If changes to the selected uncertainty factors are proposed, please identify and provide a rationale(s).**

An UF of 10 for interspecies differences and an UF 10 for intraspecies differences are warranted. An UF of 3 for subchronic to chronic is proposed by EPA; however, the alopecia endpoint occurs within 2 to 4 weeks of exposure in most species and its reversible. This was why the current assessment reduced this specific UF from the original 10 to 3. It is this reviewer's opinion that there is little or no uncertainty when alopecia occurs following oral exposure to thallium. The effect occurs in less than a subchronic time frame, and therefore the UF for the subchronic to chronic extrapolation may not be necessary for this assessment. EPA needs to counter this argument if they are concerned about toxic effects other than alopecia following chronic exposure.

**A6. Please comment specifically on the database uncertainty factor of 10 applied in the RfD derivation. Please comment on the use of the database uncertainty factor specifically for the lack of adequate developmental toxicity studies and a two-generation reproductive toxicity study, and additional uncertainty associated with the limited data available on neurotoxicity in light of the potential for neurotoxicity to represent a sensitive endpoint for thallium exposure. 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?**

An UF of 10 for deficiencies in thallium toxicity database may be deemed to be too conservative. There are indeed few well conducted oral sub-chronic or chronic studies that have targeted developmental, reproductive, neurological, and cardiovascular endpoints. However, many of the supportive IP and SC studies have examined several of these endpoints at significantly higher doses (5-50 mg/kg) than the MRI (1988) LOAEL (0.2 mg/kg) with some effects not seen at lower doses (e, g., MRI study).

***(B) Inhalation reference concentration (RfC) for thallium and compounds***
**No data are available to derive the RfC for thallium and compounds. The only published studies involving inhalation exposure include a few case reports (Hirata et al., 1998; Ludolph et al., 1986) that suggest an association between occupational exposure and toxicity (including alopecia, gastrointestinal symptoms, and neuropathy), but the route or routes of exposure in these workplace setting could not be established.**

There is simply insufficient data from which to establish a dose-response relationship, a NOAEL, LOAEL, and RfC for thallium

***B1. Has the rationale and justification for not deriving an RfC for thallium been transparently described in the document***?

The rationale and justification for not deriving an RfC for thallium has been transparently described by EPA in this document. There are currently no specific in vivo animal studies describing animal exposure to thallium by the respiratory tract. Human epidemiological studies may be confounded by oral and dermal exposures.

***(C) Carcinogenicity of thallium and compounds***
**C1. 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 thallium and compounds. Please comment on the scientific justification for the cancer weight of the evidence characterization. A quantitative cancer assessment was not derived for thallium. Has the scientific justification for not deriving a quantitative cancer assessment been transparently and objectively described?**

The human epidemiological/occupational studies are very limited, but none of them demonstrated any relationship between thallium exposure and cancer. This is supported by the lack of thallium-induced mutagenesis in relevant in vitro assays.

**George Cherian, Ph.D.**
University of Western Ontario
London, Ontario N5A 5C1
Canada

George Cherian is Professor Emeritus in the Departments of Pathology, and Pharmacology & Toxicology at the University of Western Ontario, London, Ontario, Canada. His research interests are in metal toxicology and cancer. Dr. Cherian is the author of more than 250 publications, book chapters and a book on metal toxicology. He has served in research grant review committees of MRC, Canada, NIH and US Army. He was co-chair of the WHO/IPCS health criteria document task group for zinc and a member of the task group for cadmium.

**Peer Review of EPA Toxicological Review of Thallium and Compounds**

By George Cherian

**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?*

The draft toxicological review of Thallium (Tl) compounds describes several studies on the toxic effects of Tl compounds after various routes of administration in experimental animals. Few studies on the known cases of Tl exposure in human are also presented. These are useful information on the toxicity of Tl compounds. However, I agree with the authors that most of these studies do not provide data for dose-response analysis but are useful to understand the various toxic effects of Tl and its compounds. Thus this document provides a critical evaluation of the published research on toxicity of Tl and its compounds.

The No-Observed Adverse Effect Level (NOAEL) is calculated based on a subchronic study in Sprague Dawley rats performed by Midwest Research Institute (MRI) for EPA in 1988. Three oral dose levels (0.008, 0.04 or 0.20 mg Tl/kg daily as Tl sulfate) were used in this 90 days study. This study provides only limited data but may allow to analyze a dose-response effect after oral administration of Tl compounds. Some of the limitations of this study will be discussed later in my review.

The draft document suggests a five fold lower NOAEL for Tl than the previous EPA Integrated Risk Information System (IRIS) of 1988. Both these documents are based on the same MRI subchronic study. In the previous document, the NOAEL was selected as the highest dose used (0.20 mg Tl/kg) but the present draft document wants to lower this value to 0.04 mg Tl/kg. There are several problems with this suggestion. First of all, the suggested Oral reference Dose (RfD) of 1 X 10-5 mg Tl/kg (or 0.7 ug Tl/day) for Tl is not based on any additional scientific data but a different interpretation of the same data used to derive the previous RfD. The exact reasons for the re-evaluation of the previous EPA data and different conclusion of the RfD based on two cases of alopecia in female rats are not fully explained in the draft document. Similar concerns were raised by other reviewers during the workshop held on May 19, 2008 at Arlington, VA.

The authors of the draft document try to justify the suggested low NOAEL by comparing it with other published data (page 55, para 2) but it is not a logical argument. I may disagree with the statement that the proposed LOAEL of 0.2 mg Tl/kg is *'generally consistent with other experimental studies.'* All the other studies demonstrate 3 to 9 fold level higher value in LOAEL than the proposed level, and this will make a significant difference in the derivation of RfD for Tl. Approval of this suggested level will be the lowest

level of LOAEL proposed for Tl in the literature. It may be difficult to convince the scientific community and the public that this is a valid RfD, if it is below the estimated average daily intake of the general population.

In the previous EPA document (1988), the authors did not consider these two cases of alopecia with hair follicle atrophy as a biologically significant effect, agreeing with the conclusion of the authors of the MRI study. According to Table. #4 (page 20), the incidence of alopecia is significantly increased in both male and female rats at 0.04 mg Tl/kg dose level but it was significant only in female rats at higher dose of 0.20 mg/kg. Two of the 12 female rats with alopecia also showed hair follicle atrophy, and the proposed decrease in RfD is based on these two cases. In the absence of any such significant changes in male rats with the high dose, it is difficult to conclude that alopecia, a reversible effect is a major toxic effect of Tl. Thus, alopecia does not show a dose response effect in male rats in this study. These results are surprising because male rats were more sensitive to toxicity of Tl with lower LD 50 values than female rats in acute toxicity studies (Page 52, last para).

Although alopecia is found in most of the cases of Tl exposure in both animals and human, the most significant toxic effects of Tl may be the neurotoxicity especially mixed peripheral neuropathy with paresthesia and visual disturbances. These effects were observed in both children (to treat ringworm) and industrial workers as described by a neuropathologist, Cavanagh, 1988. There are no reported experimental studies designed to investigate the neurotoxic effects of Tl using sensitive methods or behavioral changes. It is known that the biological handling of Tl and potassium is closely related because of their chemical similarity. Thus Tl can inhibit several biological functions of potassium in various organs, including the nervous tissue.

There are several limitations in the MRI study to determine LOAEL for Tl. The major weakness of this study is the lack of significant dose-response effect in male rats. As suggested by the external peer reviewers, the experimental design of the 90 days MRI study could have included one higher dose of Tl (>0.20 mg/kg), and evaluation of neurotoxic effects with changes in behavior.

There is no evidence that Tl can cause any type of cancers in human. It is not a mutagen.
There are no reported life time exposure studies to assess the carcinogenic potential of Tl. However, Tl and its compounds are very toxic to both human and animals.

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

Some of the cited references include experimental studies using more than the LD 50 values ( eg. Page 32, para 2, Leung and Ooi using 30 to 120 mg Tl/kg; page 33, Woods & Fowler, 1986 using 50 to 200

mg/kg; page 38, para 1, Lameijer & Van Zwieten, 1976; page 39, last two para, Osoro-Rico et al., 1995). These studies have little significance to the toxicity of Tl because of the high dose. Some of them reported changes in enzymes in a dying rat. My suggestion is to delete most of them from the report.

The neurotoxic effects of Tl are consistent finding in human exposure cases, and these effects should be followed up in animal studies. As stated earlier, additional experimental studies are needed to evaluate the neurotoxic effects of Tl. The observed effects of paresthesia, polyneuropathy and visual disturbance in human are very similar to the neurotoxic effects of certain metal compounds such as methyl mercury but the toxic effects may be due to different mechanisms. The neurotoxic effects of Tl may not be as prominent as that of methyl mercury because of a low half time for Tl, and it may be excreted out from the body faster than methyl mercury. The various neurotoxic effects of Tl can be tested in animal experiments with daily Tl doses below and higher than 0.20 mg/kg, and using sensitive methods (including behavioral) to evaluate neurotoxicity.

I suggest including the following references.

The human studies reported by Dr. Cavanagh's group (neuropathologist) describe some of the neuropathological effects of Tl in detail. I suggest including the following references.

Kennedy, P and Cavanagh, JB (1977) Acta Neuropathol. 39:81-88
Cavanagh, JB (1988) Recent. Ad Nerv. Syst. Toxicol. 100:177-202
A recent book chapter by G. Kazantzis in Handbook on Toxicology of Metals, edited by Nordberg, GF et. Al (2008) Chapter # 41. Thallium, p 827-37, Published by Academic Press.

The inhalation section (page 63; 5.2), should include the TLV value of 0.10 mg of Tl /m3 that has been used as OSHA standard.

3. *Please discus research that you think would be likely to increase confidence in the data base for future assessments of thallium and compounds.*

It is known that Tl and its compounds are extremely toxic. However, there is not much information about their toxicity. The published experimental studies on Tl toxicology, reviews, and this draft document suggest that one of the predominant pathological effect of Tl and its compounds is the adverse effect on the nervous system. In acute poisoning in human, the death is preceded by paresthesia, ataxia, motor weakness and convulsions. Similar effects are observed in children (treated with Tl for ringworm) and industrial workers at a dose much less than the LD50 value. Most of these effects can be considered as toxic effects affecting nervous system. There are only few studies in experimental animals, designed to study neurological effects of Tl. Kennedy and Cavanagh (1977) studied neurological effects of Tl exposure in cats after injection of 4 mg/kg Tl weekly. There are no such studies in experimental animals

after oral feeding. Therefore, it is important to undertake experimental studies to understand the initial steps involved in the development of toxic effects by feeding various dose levels of Tl compounds less than the lethal dose. In addition to markers for neurological effects, the sequel of alopecia could also be followed in these studies. These studies will provide information on the early biomarkers for Tl toxicity (alopecia or nervous system). It is important to get reliable dose-response data for alopecia and Tl in both male and female animals to find out whether there is any difference in dose response.

Another research area that needs more information is the interaction of Tl trace elements such as selenium. The interaction of Tl with potassium should be studied in detail since interference with potassium may be one of the mechanisms of Tl toxicity.

*Evaluation of the carcinogenicity of Tl*: There are no reports on detection of any type of cancers in cases of human exposure to Tl. There is also inadequate information to assess the carcinogenicity potential of Tl compounds. Animal studies with life-time exposure to tolerable dose levels of Tl are required to assess its carcinogenic potential. Such studies are expensive, time consuming and sources of funding may be difficult to find.

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 objectively described?*

There are critical deficiencies in the data base on Tl toxicity because of limited studies that can provide useful information for dose-response analysis. Therefore, several uncertainty factors have to be applied in the derivation of oral reference dose (RfD). The authors of the draft document have explained clearly all the possible uncertainty factors that should be used to derive RfD using NOAEL in the point of departure (POD) model.

Alopecia may be a critical effect of Tl compounds from both animal studies and human data. Although it is a non-specific reversible effect, one may have to use it as a marker of toxicity until better markers for Tl toxicity are available. However, the absence of statistically significant dose response effect for alopecia in male rats makes it difficult to use it as a reliable marker. Therefore one has to use several uncertainty factors. Since male rats may need higher doses of Tl for alopecia, the data for female rats may be an underestimation, and therefore, an underestimation uncertainty factor may be more appropriate. This should be explained more clearly on page 62, para 2. There may be a need to calculate separate RfD values for males and females using different uncertainty factors.

The uncertainty factor for extrapolation from animal to human is explained clearly and this factor seems to be reasonable.

The intra-human variability factor applies mainly for compounds that are metabolized to toxic or non-toxic metabolites. Since metals such as Tl are not metabolized, it is unclear whether an uncertainty factor is required for variability. There may be exceptional cases such as mercury and arsenic that are known to form alkyl derivatives which are either more or less toxic. In these cases, there is a need for uncertainty factors.

In modeling, when the uncertainty factors are high, the confidence in the RfD value will be low. The authors of the draft document has pointed out correctly that the confidence ranking for the point of departure for the calculated RfD for Tl is low. Such a low confidence level of Tl RfD may not be acceptable to scientists and the public. Thus it may be a major problem to suggest a single RfD for Tl at this time with limited scientific data. In my personal opinion, it may be better to suggest a range of RfD for Tl using various uncertainty factors because of uncertain end point, inadequate scientific data and high uncertainty factors. This approach may be more scientifically valid and acceptable to all concerned than specifically pointing out one single RfD.

5. Suggestions for other changes in the document:

1). The abbreviations do not have POD (point of departure) listed.

2). On page 41, last para, states that
*"Using a similar protocol to Hasan and Ali (1981), Hasan et.al (1977) administered --- etc."*
Something is wrong here. How can a previous study (1977) use a protocol used in a later study (1981)?. This section needs re-writing.

3). Several of the typographic errors make the report difficult to read. The report should be proof read to correct all the errors, including double typing.

**Chemical-Specific Charge Questions:**

A. Oral reference dose (RfD) for Tl.

A1. There are only few studies that provide toxicological data after oral feeding of various doses of Tl compounds. Of these, the experimental design of MRI, 1988 study may be more appropriate for derivation of RfD . Thus, its selection is justified. However, few other studies could also be considered as discussed by others in the workshop.

A2. Alopecia is a common critical effect of Tl exposure in both animals and human. Therefore, its use as a biomarker is justified until a better marker for adverse effect such neurotoxic effects is identified. There are few drawbacks to the use of alopecia as an endpoint for RfD derivation. Alopecia is a reversible effect, and the MRI, 1988 study does not show a statistically significant dose response effect for alopecia in male rats.

A3. In a previous document in 1988, EPA used the same MRI study for evaluation of NOAEL for Tl, and selected the dose of 0.25 mg/kg-day Tl sulfate as the NOAEL. But in the present draft document, the NOAEL value has been lowered to 0.05 mg /kg. I cannot fully agree with this approach because it is based on a change in interpretation of the data rather than any new scientific evidence. There is no additional scientific data on alopecia as a marker for Tl exposure since 1988. Moreover, the male rats do not show a statistically significant dose response for alopecia, and that was the basis for a higher NOAEL in previous document.

A4. The traditional NOAEL- LOAEL approach to define RfD is ok, if there is a well defined end point for adverse effect and sufficient statistically significant data without many variables. The present analysis in the draft document has suggested a single RfD value using a large number of uncertainty factors. The confidence level of the proposed RfD is low because of inadequate scientific data and undefined end point for adverse effect. I suggest deriving a range of RfD values rather than a single RfD value.

A5. The uncertainty factors are needed for derivation of RfD because of lack of statistically significant data with several variables. Most of the proposed uncertainty factors are needed except the factor for intra-human for Tl. This factor is required for toxic compounds that undergo metabolism to form either toxic or non-toxic metabolites. Since Tl is not metabolized, the sensitivity to Tl in human will be insignificant, and this uncertainty factor may not be needed.

A6. The selection of the uncertainty factors for derivation of RfD is explained in detail in the draft document. These factors are required because of lack of data and variability of available data on toxicity of Tl. I am proposing to change the single RfD to a range of RfD for Tl to make this a scientifically valid exercise.

B1. There is no data available to derive an inhalation reference concentration RfC. However, OSHA standard of TLV 0.10 mg of Tl/m3 is available, and this should be reported in the draft document.

C1. A life time oral exposure study for Tl in animals is lacking, and therefore, the carcinogenic potential of Tl is not known.

The external reviewers of MRI (1988) study and public comments (submitted) also have pointed out many of the concerns that were discussed here.

**Summary and Conclusions:**

1. This draft document provides a comprehensive toxicological review of published data on Thallium and compounds. Some of these publications are critically analyzed.

2. There are only few studies on oral administration of Tl. A review of these studies show that most of them have used either high doses or the study period was short. In one study (Downs, et.al., 1960), both low and high doses were used but the analysis of the data was difficult because of mortality in both control and treated groups. So the selection of an unpublished study by MRI (1988), and sponsored by EPA for derivation of RfD for Tl is justified.

3. Alopecia is a critical effect of Tl exposure in both animals and human, and is used as an end point to determine NOAEL & LOAEL, and derive RfD for Tl in this draft document. The selection of alopecia is controversial but is justified because none of the studies have identified markers for neurotoxicity, except one study in cat after injection of Tl compounds. The adverse effects on nervous system are frequent findings in human exposure cases of Tl.

4. The previous NOAEL value estimated by EPA was 0.20 mg Tl/ kg, the highest dose used in the MRI (1988) study. This value was selected because the authors of the study did not find any biologically significant alopecia or any other toxic effects. But using the same data, the present draft document has identified 0.04 mg Tl/kg as NOAEL and 0.2 mg Tl/kg as LOAEL. These values were based on a different interpretation of the same data but not based on any new scientific finding.

5. A number of uncertainty factors are used for the calculation of RfD for Tl in the draft document, and most of them are justified because of undefined end point and inadequate statistically significant data base. However, the intra-human uncertainty factor may not be needed for Tl.

6. The suggested RfD value (0.7 ug Tl/day) is much lower than that proposed by WHO/IPCS health criteria document (10 ug Tl/day). It is also less than the estimated daily intake of Tl (2ug Tl/day) in diet in general population. Because of these and the low confidence level of the calculated single RfD value, this suggestion should be reconsidered. I suggest to calculate a range of RfD values using various uncertainty factors and two NOAEL estimates. This approach may be more acceptable to the public and scientific community rather than a single RfD value for Tl.

**Lucio G. Costa, Ph.D.**
University of Washington
Seattle, WA 98105

Lucio G. Costa is Professor of Toxicology in the Department of Environmental and Occupational Health Sciences at the University of Washington and Professor of Pharmacology at the University of Parma Medical School. His research interests are in the field of neurotoxicology, particularly in the area of developmental effects and mechanisms of pesticides, alcohol, metals and other environmental pollutants. He has also an active research program in the area of gene-environment interactions, as they relate to susceptibility to pesticides and neurodegenerative diseases. Dr. Costa is the author of more than 300 publications, book chapters and books in the field of toxicology, and is a member of numerous state, national and international panels and committees dealing with toxicology and risk assessment.

**Review of "Toxicological Review of Thallium and Compounds"**

**General Charge Questions:**

**1.** Overall, the Toxicological Review is logical, clear and concise, and EPA has accurately synthesized the available information. However, the 1988 MRI study, which is the basis for the derivation of the RfD could have been better described and commented on. For example, there is no discussion on the incidence of miosis which, though reversible, may be an indicator of neurological effects. Miosis was present (control, vehicle, 0.008, 0.04, 0.2 mg Tl/kg/day) in 0, 0, 5, 7, 15 male rats and 2 ,0, 1, 11, 8 female rats. The same considerations also apply to other end-points of the MRI study. Other in vivo studies, which may be relevant to the overall assessment, are discussed (and dismissed) somewhat superficially, and a more in depth discussion is recommended. More detailed information on concentrations used in some in vitro studies (p. 430) would also be useful. Furthermore, some reorganization may be useful. For example, on p. 38, section 4.4.3, line1-3, it is stated that no studies of thallium neurotoxicity following exposure by the oral route were identified. Yet, on p. 20, line 1, the study by Manzo et al. (1983) is described, in which thallium was administered by the oral route, and which focuses primarily on neurotoxicity end-points.
Finally, the title (Toxicological Review of Thallium and Compounds) is a bit unclear and confusing. Potential alternatives: TR of thallium and its salts or TR of thallium and thallium compounds.

**2.** This reviewer is not aware of any additional study that should be considered. However, no detailed literature search was carried out. A rapid Medline search identified several articles dealing with the toxicity of thallium that were not cited in the documents. Though these were not reviewed to determine whether new important additional information was present, there appears to be a number of studies that the document may want to include.

**3.** Additional information is needed on various aspects of thallium toxicity. First, a developmental toxicity study, and a neurotoxicity study, carried out according to established guidelines, would be useful to increase the limited information available in these areas, and to better define the additional UF of 10 that was used in the present assessment to compensate for an incomplete database. Second, a chronic bioassay would be useful at two levels: first, it will provide information on chronic non-cancer effects, thus addressing the current UF of 3 used for subchronic to chronic extrapolation; second, it will provide information on cancer end-points that would be necessary for deriving carcinogenic potential. Also, mechanistic studies on thallium-induced alopecia and neurotoxicity would be useful to increase confidence in overall risk assessment.

**4.** The sources of uncertainty have been adequately addressed. In addition to the standard UFs of 10 for species extrapolation and intraspecies variability, two additional UFs were used: one (UF=3) to take into account the subchronic (90 day) nature of the main available study; the other (UF=10) to account for an inadequate database, since solid and reliable data on developmental, reproductive and neurotoxicity are not available. This reviewer also agrees on the levels of confidence that are described in the document (p. 67).

**Chemical-Specific Charge Questions:**

**A1.** The 90 day oral gavage study by MRI (1988) appears to be the only study in which thallium was given by the oral route, for an extended period of time, and at multiple dose levels. In addition, the study was conducted under established guidelines, in GLP, and was previously peer reviewed. The study is described in the document on pp. 17-20. Major end-points are described. However, it would be useful to report data on exophthalmos and lacrimation in tabular form. Furthermore, there is no mention of miosis, which was reported in the MRI study, and which was noted as a potentially relevant end-point by the reviewers of the study. This reviewer is not aware of any other study that could have been used as a primary study to derive the POD and the RfD. However, given the paucity of data available, other studies could be analyzed more in depth and information compared to that obtained from the MRI study.

**A2.** Alopecia is a characteristic effect of thallium exposure. Several human reports have indicated that the presence of alopecia, in the presence of less well defined gastrointestinal and neurological abnormalities, is strongly suggestive of thallium poisoning. Alopecia is a reversible phenomenon; however, this reviewer would agree that it represents an adverse effect. The mechanism of thallium-induced alopecia is not known. Hypotheses relate to the ability of thallium to interfere with incorporation of cys(e)ine into proteins, its ability to interact with riboflavin, or to a general depression of mitotic rate resulting possibly from metabolic and energetic impairment. Direct toxicity of thallium on hair follicles may also be involved. The document could provide a more detailed discussion of mechanisms of alopecia, particularly with regard to thallium.

**A3.** In the key MRI (1988) study, alopecia was increased in both male and female rats treated with thallium. Such increase was statistically significant in males and females exposed to 0.04 mg Tl/kg/day, and in females exposed to 0.2 mg Tl/kg/day, as determined by EPA's analysis (Fischer's exact test). Of the cases of alopecia, most were attributed by the MRI study authors to barbering behavior. Those cases not attributable to barbering behavior are shown in Table 1. It is not known whether differences observed in treated animal (particularly females) are statistically significant. However, a clear dose response is seen

in female rats. Hair follicle atrophy was only measured in vehicle controls and in the high dose thallium group, and was found in two females (out of 20, of which 12 with alopecia and 5 with alopecia not attributed to barbering behavior).

Table 1. Incidence of alopecia not attributed to barbering behavior

| Dose (mg Tl/kg/day) | Males | Females |
|---|---|---|
| 0 (untreated controls) | 1 | 0 |
| 0 (vehicle controls) | 0 | 0 |
| 0.008 | 2 | 1 |
| 0.04 | 4 | 3 |
| 0.2 | 1 | 5 |

Since alopecia is a key characteristic effect of thallium exposure in humans, it would appear that the choice of alopecia (not attributable to barbering behavior, and supported by histological evidence of hair follicle atrophy) should be considered as an adverse effect. This reviewer would support the choice of the high dose of thallium (0.2 mg/kg/day) as a LOAEL, rather than a NOAEL. However, statistical analysis of data summarized above in Table 1 is recommended. The data presented above in Table 1 would be amenable for a BMD approach. While data on hair follicle atrophy are available only for the control and high dose thallium groups, information on alopecia is available for all treatment groups. In this respect, the middle dose of thallium (0.04 mg/kg) may become a LOAEL rather than a NOAEL, if alopecia itself is considered as an adverse effect. Most importantly, however, several other end-points measured in the MRI study which show clear dose-response effects, should be considered for a more comprehensive hazard evaluation.

**A4.** This reviewer would support the use of a NOAEL/LOAEL approach to determine the POD. This is adequately described in the document. Most likely, a BMD approach would not be feasible given the lack of data on the key endpoint (hair follicle atrophy) at each dose level. However, as indicated, a BMD approach can be applied to the data on alopecia as well as other important parameters (e.g. exophthalmos, lacrimation, blood chemistry etc.).

**A5.** The overall UF applied to the POD is 3000. This includes two UFs of 10 each for species extrapolation and intraspecies differences, an UF of 3, since the main study (MRI, 1988) is a subchronic study, and an UF of 10 for insufficient database (particularly information on developmental and reproductive effects and on neurotoxicity). The choice of UFs is adequately described. The overall UF of

3000 does not differ from that of a previous evaluation; however, the UF for subchronic to chronic extrapolation has been lowered to 3 (from 10), while the UF for insufficient database has been increased to 10 (from 3). This reviewer agrees with the EPA's rationale for this choice. Two or three additional studies (as indicated earlier) would be needed to eliminate these two additional UF.

**A6.** The absence of specific dose-response studies addressing the issues of neurotoxicity, reproductive toxicity and developmental toxicity of thallium, suggest that an UF to account for the incomplete database is appropriate. This is supported by evidence that neurotoxicity is seen in humans upon (high) exposure to thallium and was also seen in animal studies (e.g. Manzo et al. 1983). Reproductive toxicity was also seen in animal studies. Of note is a study by Wei (1987) on reproductive effects in male mice exposed to thallium carbonate in drinking water for six months. This study, published in a Chinese journal and summarized on p. 25 of the document, indicates that a 0.001 mg/L concentration of thallium is a LOAEL for male reproductive toxicity. Assuming a daily water consumption of 5 ml/day and a body weight of 30 g during treatment, this will provide a dose of 0.00017 mg/kg/day as a LOAEL, three orders of magnitude below that derived from the MRI (1988) study. Perhaps the indicated limitations of this study should be discussed more thoroughly. The limited information available does not suggest that developing animals may be more sensitive to thallium toxicity.

**B1.** Since the only available data on inhalation exposure to thallium were a few case reports, no RfC could be derived. This is properly explained and justified.

**C1.** Given the absence of any information (cancer bioassay, epidemiological data) on the carcinogenicity of thallium, no quantitative cancer assessment was derived. This is properly explained and justified.

Overall, the main problem with attempting to derive an RfD for thallium is the limitation of the database. If no new study becomes available, the high default UF of 3000 should be applied to the POD value. This may underscore more the uncertainty of the database than the actual risk for humans from exposure to thallium. It could be, however, possible to reevaluate some studies mentioned in the document, and particularly the 1998 MRI study, and consider more thoroughly some additional end-points other than alopecia (with follicular atrophy). BMD calculations can be done for several of these end-points. Such approach would somewhat strengthen the overall hazard evaluation of thallium. However, given the data deficiencies, the confidence in the overall RfD would remain low.

**George Daston, Ph.D.**
Miami Valley Laboratories
Proctor and Gamble Company
Cincinnati, OH 45253

Dr. Daston received his B.S., magna cum laude, from the University of Miami (biology; chemistry/geology). He received his Ph.D. in developmental biology and teratology also from the University of Miami. In his extensive career as a research toxicologist, Dr. Daston has made significant contributions in the areas of developmental biology, teratology and toxicology, especially mechanisms of normal and abnormal development, nutrient-toxicant interactions, in vitro alternatives in teratology and toxicology, functional teratology, fluid balance in development, and risk assessment. He served as a National Research Council Research Associate in the Developmental Biology Division of EPA's Health Effects Laboratory upon receiving his Ph.D. He is also an adjunct associate professor of pediatrics at the Children's Hospital Research Foundation at the University of Cincinnati. Dr. Daston has served on EPA's Environmental Health Committee as a consultant on the Reproductive Toxicity Risk Assessment; the Risk Assessment Task Force of the Society of Toxicology; and Chair of the Reproductive and Developmental Effects Subcommittee for the American Industrial Health Council. He was a member of the Advisory Group on Birth Defects and Developmental Disabilities Program for the Centers for Disease Control and Prevention. Dr. Daston has given dozens of invited lectures before numerous scientific bodies; authored or coauthored well over 100 published scientific articles; and edited or co-written dozens of chapters in biology and toxicology texts. Since 2003, he has been Editor-in-Chief of *Birth Defects Research B: Developmental and Reproductive Toxicology;* and on the editorial board of *Human and Ecological Risk Assessment*.

**Peer Review of Toxicological Review of Thallium and Compounds**

George Daston
May 1, 2008

**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?

I found the review to be presented in a logical format. It was easy to read. Information from accidental or intentional poisonings, epidemiology studies, and animal studies were synthesized in an objective way, with the purpose of demonstrating which effects associated with thallium exposure are biologically plausible. Despite the fact that thallium exposures have occurred in humans, the toxicology literature is spotty and incomplete. The centerpiece of the animal toxicology data is a single, well-conducted subchronic toxicity study in rats. Although there are other animal studies in the literature, they have limitations; some in study design, some in statistical power or lack of multiple dose groups, some in study quality. EPA has tried to make the most of the subchronic study it has, including subjecting it to expert peer review. Data on reproductive and developmental effects are limited. So are neurotoxicity evaluations, despite the fact that neurological symptoms appear to be one of the hallmarks of acute thallium toxicity in humans. There are some interesting studies on possible mechanisms of thallium toxicity, but nothing definitive. There are no animal studies on the chronic effects of thallium. Therefore, it is not possible to draw conclusions about cancer hazard.

The epidemiology data are also limited. It is difficult to put much confidence in them. The case reports of human poisonings are useful in identifying a pattern of symptoms associated with high-dose exposures, but accurate dose estimation from these reports is not possible. Despite these considerable shortcomings in the overall database for thallium (which are acknowledged in the report), the authors of the review have done a good job at synthesizing the available information and presenting it in a logical fashion.

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

I am unaware of other studies that would improve the assessment.

3. Please discuss research that you think would be likely to increase confidence in the database for future assessment of thallium and compounds.

The database on thallium is sufficiently spotty that it is at the hairy edge of being adequate for an IRIS assessment. This is tacitly acknowledged by the application of an aggregate 3000 x uncertainty factor to the point-of-departure (POD), beyond which it is the policy not to calculate a reference dose because the

uncertainty is too great. Therefore, there are a large number of studies that would be useful in increasing the confidence of the assessment. In my opinion, the most critical are:

An up-to-date subchronic study that includes neurotoxicity endpoints would be helpful in resolving some of the outstanding questions about subchronic toxicity. It would be possible in such a study to more thoroughly evaluate the basis for alopecia, which appears to be a plausible effect of thallium, but for which the interpretation is complicated in the existing subchronic study by a high background incidence and distinctions about animal behavior that are not commonly made in such studies. Additional clinical chemistry, functional, and histopathological assessments would be useful in determining the source and intensity of the clinical chemistry changes observed in the MRI subchronic study.
A reproductive toxicity study conducted according to an established regulatory guideline would improve the ability to interpret the results in Kunming mice observed by Wei.

Studies on metabolism of thallium that address the question as to whether the body has the capability to convert thallium from one valence state to the other would be important in determining the extent to which studies on monovalent Tl are relevant to the toxicity of trivalent Tl. If there is little or no conversion, then it will be important to characterize the toxicity of trivalent Tl separately.

A modern in vivo genotoxicity evaluation would be useful in improving our ability to conclude whether Tl is genotoxic. Metals are notorious for producing false positive results in vitro. The sole in vivo study on Tl, a dominant lethal study, had interpretation problems and cannot be relied on to address the question. Depending on the outcome of a genotoxicity assessment, the value of conducting a lifetime chronic bioassay could be weighed.

More studies on the potential mechanisms of action of Tl would be very useful, and perhaps the most economical way of addressing uncertainties about the target organs for Tl and the plausibility of some of the responses that have been observed.

Pharmacokinetic studies that would support the development of a PBPK model for thallium would be useful in replacing the interspecies uncertainty default.

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?

I found the discussion of uncertainty to be comprehensive. It would have been helpful, however, to have this discussion all in one place.

*Post-meeting comment: As noted in my response to question 3, the database on thallium toxicity is so incomplete that it is a borderline call as to whether it is sufficient to support an IRIS assessment. The aggregate uncertainty factor of 3000, which is the maximum value allowed under EPA guidance for risk assessment, can be interpreted as an acknowledgement of the weakness of the data set. Therefore, it is important to openly discuss in the document the pros and cons of calculating an IRIS RfD for thallium.*

**Chemical-Specific Charge Questions:**

(A) Oral reference dose (RfD) for thallium

A1. The 90-day oral gavage study by Midwest Research Institute (MRI, 1988) was selected as the basis for the RfD. Please comment on whether the selection of this study as the principal study has been scientifically justified. Has the 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.

I believe that the selection of the MRI study was scientifically justified. It is one of the few studies on Tl that was conducted according to a testing protocol that used a relevant route of exposure, multiple dose groups, and enough animals per group to provide adequate statistical power to draw conclusions. EPA arranged an external peer review of the study, which confirmed that the study quality was satisfactory. The study is described sufficiently well in the document to allow the reader to understand what was done in the study, and what was found. Inclusion of the external peer review comments was also helpful in supporting the interpretation of the study.

There were one or two other studies that were considered for use as the principal study, but the document provides good rationale as to why they were not as good choices as the MRI study.

> A2. Alopecia (hair loss) was selected as the most appropriate critical effect for the RfD. EPA characterizes alopecia as an adverse effect. Please comment on whether the science and mode of action information supports alopecia as an adverse effect… Please comment on whether the selection of this critical effect has been scientifically justified. Is EPA's choice transparently and objectively described in the document? Please provide a detailed explanation. Please identify and provide the rationale for any other endpoints that should be used instead of alopecia to develop the RfD.

I disagree with the use of alopecia as the critical effect for RfD determination for Tl. My opinion is based less on the question as to whether alopecia is an adverse effect, but that it was not clear to me from either the description of the study or the peer review as to whether this effect was really treatment-related.

As to whether alopecia is an adverse effect, I would consider unwanted hair loss as undesirable. If it were definitive that the alopecia seen in the MRI study were treatment-related, and if we knew whether the dermal histological effects associated with alopecia were in fact on a continuum to more significant dermal effects, then I would have no quarrel as to the choice of alopecia as an adverse effect. However, neither of these is known. (I do believe that the human poisoning case reports support the biological plausibility of alopecia as a Tl-related effect.)

As to whether the effect was treatment-related, I find it difficult to conclude definitively that it was. There was a high background incidence of alopecia in the study population. There was also the observation in the study of "barbering behavior" in the animals that may have accounted for much of the alopecia. There were only two animals in the high dose group for which the alopecia could not be attributed to this behavior. My own experience with lab animals is that there can be any number of reasons for hair loss, including changes in caging or husbandry. The fact that there was such a high background incidence of alopecia suggests to me that the effect could have been unrelated to treatment.

I believe that the review's authors have made a reasonable case for alopecia as the critical effect, but not a compelling one. There were other observations at the same dose level, including significant changes in clinical chemistry and exophthalmos, that I believe would be easier to conclude as being Tl-related. An RfD based on one of these as the critical effect would be more defensible, and would be protective against alopecia.

> A3. …Is EPA's interpretation of the study findings scientifically justified? Has the interpretation of the findings been transparently and objectively described in the document? …Please comment on whether EPA chose the appropriate data set and the appropriate statistical test for this analysis. …Is the study author's conclusion that the high dose (0.25 mg/kg/day thallium sulfate) represents a NOAEL justified and supported by the study data?

As noted in the response to the previous question, I find it difficult to conclude definitively that the alopecia in two animals, which could not be attributed to either barbering behavior or normal cyclical hair

growth, was in fact treatment-related. The main support that they were is the observation of atrophy of hair follicles in those two animals. However, hair follicles go through a natural cycle of activity and inactivity; it is not clear to me from the descriptions of atrophy that what was being observed was anything more than an observation of a normal condition in these animals. There was no systematic evaluation of dermal tissue from other areas in the same animals, nor was there an evaluation of hair follicle status in all dose groups. Furthermore, this is not a standard assessment in subchronic studies, so it is not possible to know whether what was observed is within the range of normal or not.

As to whether Fisher's exact test was the right test to use to evaluate quantal data in separate dose groups, I believe that it was. However, the choice to include all animals that had alopecia is problematic, especially given that the study investigators had found non-treatment-related reasons for most cases.

As noted in the response to the previous question, the conclusion as to whether the observation of alopecia is an adverse one is a matter of judgment. The investigators who wrote the original study report have a different interpretation of the alopecia. It is less clear to me why they concluded that the clinical chemistry observations or exophthalmos were not adverse.

*Post-meeting comment: Given the other endpoints were affected at the same dose levels that produced alopecia, and that some of these endpoints could be considered to be adverse, it would be worthwhile to model the dose-response relationships for these to determine whether a benchmark dose for one or more of these would be a better point-of-departure for risk assessment. The effects that should be modeled are lacrimation, exophthalmos, and the significant clinical chemistry findings.*

A4.

Assuming that alopecia remains the critical effect, then there is no other choice than to use a NOAEL as the POD. The data are not amenable to the application of BMD methods. On the other hand, if another critical effect is used, e.g., the change in one or more clinical chemistry parameters above the range of normal, then it may be possible to model the dose-response relationship and calculate a benchmark dose.

A5.

I believe the choice of uncertainty factors to be justified and to be consistent with EPA risk assessment policy. There is no basis to move away from default values for inter- or intraspecies considerations. Because the critical effect is from a subchronic study, a factor of 3 to account for potential chronic exposure is appropriate. Because the data base is so spotty, there needs to be a way to account for database insufficiency. Whether a 10 x factor is justified for this is debatable, but this is a judgment call that seems reasonable to me.

A6.

I believe that the sources of database insufficiency have been well documented in the review. Neurotoxic effects are among the hallmark effects of thallium poisoning, and is neurotoxicity has not been adequately evaluated in the animal studies that have been run. There is an indication of an effect on sperm production in a mouse study. This study was not adequate to identify a NOAEL for the effect. It is not likely that the NOAEL for this effect would be 10 x below that for the observations seen in the subchronic study, but this cannot be concluded with certainty. The developmental toxicity studies that have been done, although not state-of-the-art, appear adequate to conclude that developmental toxicity would not drive the risk assessment for Tl.

B1.

The rationale for not conducting an inhalation risk assessment is clear and transparent. Although observations from occupational studies suggest that Tl can have effects by the inhalation route, the studies are not good enough either in dosimetry or excluding confounders to be used as the basis for risk assessment. There are no animal studies that are useful as the basis for RfC determination. The PK data for Tl are insufficient to support route-to-route extrapolation.

C1.

I agree that there is no basis for conducting a cancer assessment on Tl. There are no chronic animal studies, and there are no useful epidemiology studies. The genotoxicity data on Tl is inconclusive. There is no way to provide a cancer assessment on Tl.

**Robert Hoffman, M.D.**
New York University
School of Medicine
New York, NY 10016

Robert S. Hoffman received a BA in chemistry from Brandeis University in 1980 and immediately entered New York University School of Medicine. Following earning his MD, he completed a 3-year internship and residency in Internal Medicine also at the New York University School of Medicine, followed by a Fellowship in Medical Toxicology. He achieved Board Certification in Internal Medicine, Medical Toxicology, and Emergency Medicine. In 1989 Dr. Hoffman became the director of the Fellowship in Medical Toxicology at the New York City Poison Center, and in 1994 he became the Director of the New York City Poison Center.

Dr. Hoffman has authored over 200 peer-reviewed publications in various aspects of toxicology that include basic science, animal and clinical investigations. He has also authored numerous textbook chapters for major references in Medicine and Emergency Medicine. He lectures around the world on various aspects of toxicology and has helped to establish poison control centers in both Europe and Asia. Dr. Hoffman has held offices in 3 American Toxicology Societies, including a member of the board of trustees of the American Academy of Clinical Toxicology; a member of the board of directors of the American Association of Poison Control Centers; and Secretary/Treasurer, Vice President, and President of the American College of Medical Toxicology. Dr. Hoffman also is currently on the editorial board of Toxicological Reviews and is a member of the senior editorial board of Clinical Toxicology.

Dr. Hoffman began to write for Goldfranks Toxicological Emergencies in the 4th edition and has been an editor since the 5th edition. Along with the current 8th edition, Dr. Hoffman has assumed the lead responsibility for coordinating the first edition of the companion handbook.

With regard to thallium, Dr. Hoffman has been involved in direct bedside care of thallium poisoned patients, international consulting on cases of malicious thallium poisoning, basic science laboratory investigation, and animal research.

**Peer Review Workshop for Toxicological Review of Thallium and Compounds**

**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?*

The document reviews most of the English literature with regard to acute, chronic and sub-toxic doses of thallium. Unfortunately several environmental surveys and other sources of data where not included. Most strikingly is a lack of attempt to seek out foreign (not translated) material and incorporate this material into the review. Thallium is not strictly a problem of the English speaking world and I must be critical of any attempt that failed to consider sources from other languages (I will be specific below). I find it hard to believe that this is the first attempt (worldwide) to set standards for thallium. The failure to include information from other countries or comment as to whether thallium standards have been set in other countries is a major weakness of the review. The synthesis is a somewhat less than ideal in that it is interjected with editorial comments (some from original authors). It might have been more desirable to present the facts and leave the interpretation to the reader. Numerous typographical errors and poorly formed sentences distract from the actual content of the review.

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

Granero S, Domingo JL. Levels of metals in soils of Alcalá de Henares, Spain: human health risks. Environ Int. 2002 Jul;28(3):159-64.

The foreign experience may be very significant. These are papers that I was easily able to find. They should be obtained, translated by a reliable source and included in the review. I am certain that many other (non-English) documents exist and should be considered given the very limited amount of material available here.

A. Kamil'dzhanov AKh. [Experimental substantiation of maximum permissible concentration of thallium carbonate in environmental air]. Gig Sanit. 1993;(5):8-10. PMID: 8063180
B. Gerasimova IL. [Establishment of MPEL for thallium iodide activated cesium iodide in the working zone air]. Gig Tr Prof Zabol. 1991;(1):31. PMID: 2060817
C. Viereck L, Kramer M, Eikmann T, König W, Bertges WD, Gableske R, Krieger T, Michels S, Exner M, Weber H. [Determining guidelines for metals in children's playgrounds in North Rhine-Westphalia]. Offentl Gesundheitswes. 1990;53(1):7-15. PMID: 2150550

D. Krasovskiĭ GN, Kenesariev UI. [Methodological outline for the experimental substantiation of a system of indices of the adverse effect of metals on the health status of the population (the example of thallium)]. Gig Sanit. 1984;(2):22-5. PMID: 6714685
E. Zasukhina GD, Vasil'eva IM, Sdirkova NI. [Approach to the determination of the mutagenic potential of environmental pollutants with the example of detecting the mutagenic action of thallium carbonate]. Dokl Akad Nauk SSSR. 1980;250(3):766-8.PMID: 7353468

*3. Please discuss research that you think would be likely to increase confidence in the database for future assessments of thallium and compounds.*

The entire analysis is based on determining a NOAEL or LOAEL dose from very limited literature. Using an excepted animal model of chronic exposure and a well designed, blinded, controlled and properly powered study would greatly increase confidence. This study should look at a wide variety of endpoints and be sufficiently robust to calculate a dose response curve for each. Ultimately it should undergo a formal peer review process in a major medical journal. Short of this, only a much larger and very costly human investigation of an exposed population would suffice. While both would be of interest, the animal model might be more practical.

*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?*

Section 5.1 deals with the standard applications of uncertainty to the proposed reference dose. For the majority of decisions the default values are selected. The large uncertainty value of 3000 is reflective of the very limited data available for analysis. The presented data are so poor that this review questions whether such an analysis is even valid.

Whereas some limited dose response data exist for human dermal exposure, for example it was routine practice of dermal administration of thallium salts to children with ringworm yielded valuable dose-response data. A closer look at papers from Munch in the 1930's (one cited in the reference list of the review) nicely demonstrates the distinction between epilation and toxicity and suggests a possible range for which animal effects can be correlated with human effects. It is entirely unclear whether the addition of these data would improve uncertainty.

The very limited data set can be further appreciated by evaluating the proposed value of $1 \times 10^{-5}$ mg/kg/day (10 ng/kg/day). This suggests that a 70 kg adult can consume 700 ng/day (0.7 mcg/day). The $90^{th}$ percentile from NHANES data for adults lists a urinary elimination of about 0.380-0.390 mcg/L

which is about 0.760-0.780 mcg/day in the urine alone. Since thallium has substantial fecal elimination, this suggests that if the standard is adopted, greater than 10% of Americans, and in reality, probably close to 50% of Americans would ingest more than the daily reference dose. Because there is no evidence that the thallium in the current US diet poses any threat and because there is no possible remediation, even if this calculation were correct, adoption of this standard would produce unnecessary concern.

As will be discussed below, a more thorough use of the 1988 MRI study would probably further lower the RfD suggesting possibly that the entire US population is exposed above threshold. While this might be acceptable to some, this reviewer views such an analysis as an example of how poor the existing data are and further questions the validity of the analysis.

**Chemical-Specific Charge Questions:**

*(A) Oral reference dose (RfD) for thallium*

*A1. The 90-day oral gavage study by Midwest Research Institute (MRI, 1988) was selected as the basis for the RfD. 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.*

There are significant concerns about this study. Neither the choice of animal model, the doses given, nor the duration of investigation are well substantiated. There is no presentation of sample size, statistical analysis or power calculation. These concerns are compounded by the lack of publication in a peer-reviewed journal so that the details of the methodology and results have not been critically reviewed. I am not certain that there are better studies but the external peer reviewers highlight some of the essential weaknesses of the analysis. Furthermore, the analysis of the study is incomplete and somewhat misleading in that it chooses to discount many finding that might be of value (see below)

*A2. Alopecia (hair loss) was selected as the most appropriate critical effect for the RfD. EPA characterizes alopecia as being an adverse effect. Please comment on whether the science and mode of action information supports alopecia as an adverse effect. EPA has stated: "Whether alopecia is itself an adverse effect merits consideration. In humans, alopecia is generally reversible upon cessation of thallium exposure. Alopecia, however, appears to be a part of a continuum of dermal changes observed following thallium exposure, as well as one of a spectrum of effects on target organs that include the nervous and gastrointestinal systems. For these reasons, alopecia supported by two cases of hair follicle atrophy is considered an adverse effect." Please comment on whether the selection of this critical effect has been scientifically justified. Is EPA's choice transparently and objectively described in the document? Please provide a detailed explanation. Please identify and provide the rationale for any other endpoints that should be used instead of alopecia to develop the RfD.*

This reviewer agrees with the EPA's opinion that alopecia is in fact an adverse effect. Thallium is clearly toxic to hair follicles and is repeatedly discussed in all clinical cases of poisoning. In one series of

patients, alopecia was the only finding that was universally present. Additionally, there is little doubt from the "therapeutic" use of thallium salts that alopecia is an expected effect that is dose related and precedes other serious manifestations of toxicity. While there is no evidence that alopecia in and of itself has any lasting harmful effects, it serves as one simple biomarker of toxic exposure. In all cases including the most severely poisoned individuals, alopecia is reversible. However, that being said this reviewer disagrees with the choice of alopecia as the 1988 MRI study includes several findings (biochemical parameters, exophthalmos, miosis and changes in coat) that appear to show better dose-response effects than alopecia. Furthermore, there was no attempt to study neuropathy in these animals which may have also provided a useful marker of toxicity.

*A3. At the high dose in the MRI (1988) study, two female rats exhibited moderate to severe alopecia that could not be attributed to self-barbering or normal cyclic hair growth. Histologic examination of skin samples from these high-dose females showed atrophy of hair follicles. EPA considered these findings to be adverse, and thus the high dose in this study (0.25 mg/kg-day thallium sulfate) to be the lowest-observed-adverse-effect level(LOAEL). The mid-dose group (0.05 mg/kg-day thallium sulfate) was identified as the no-observed-adverse-effect level (NOAEL). Is EPA's interpretation of the study findings scientifically justified? Has this interpretation of the findings been transparently and objectively described in the document?*

The use of these findings is questionable. The mid-dose group also had alopecia and although it may be a result of barbering, barbering is not a normal behavior and suggests that the animals are under stress, possibly in pain, thus this can not be an NOAEL. It is not valid to discount a sex difference as clearly in other models sex may affect sensitivity to a given toxin. The entire analysis seems flawed, or questionable at best.

*As part of the evaluation of alopecia as a critical effect for the RfD, EPA performed a series of Fisher's Exact Tests to determine if the incidence of alopecia in any of the three dose groups was statistically significantly elevated above controls using all cases of alopecia reported by MRI (1988). Please comment on whether EPA chose the appropriate data set and the appropriate statistical test for this analysis.*

It is unclear to this reviewer whether this is the correct scientific approach. A Fisher's exact test is specifically designed for a simple comparison. In a dose response relationship generally a more complex analysis is performed. It is quite likely that there would be no significance had that approach been taken. I would suggest that the original data be transmitted to a statistician for independent review. There are other parameter (stated above) that show a dose-response and therefore might be amenable to more of a benchmarking dose analysis.

*The study investigators reached a different interpretation of the study findings than did EPA. The investigators considered alopecia to be attributable to the cyclic pattern of hair growth in rodents and, consequently, did not consider these findings to be biologically significant. The high dose (0.25 mg/kg-day thallium sulfate) was identified in MRI (1988) as the NOAEL. Is the study authors' conclusion that the high dose (0.25 mg/kg-day thallium sulfate) represents a NOAEL justified and supported by the study data?*

Clearly one of the many weaknesses of the 1988 thallium study is the lack of blinded interpretation of the results. While the investigators attribute the findings to barbering, the presence of follicle changes suggest that this is truly a toxic effect. This reviewer disagrees with the EPA interpretation of this dose as an effect possibly near the LOAEL and not an NOAEL. Barbering represents stress and possibly pain. The study's own investigators' interpretation of this data adds considerable doubt on their ability to interpret the other findings in their study and further calls into question the choice of this study as the critical data for determining the reference dose.

*A4. The traditional NOAEL-LOAEL approach was used to define the point of departure (POD) for the RfD. A benchmark dose (BMD) analysis was considered but was not conducted because of the nature of the data set for alopecia. Please provide comments with regards to whether a NOAEL-LOAEL approach is the best approach for determining the POD. Has the approach been scientifically justified? Is it transparently and objectively described? Please identify and provide a rationale for any alternative approaches for the determination of the POD, and if such approaches are preferred to EPA's approach.*

If one accepts that alopecia represents and affect than I agree that the NOAEL-LOAEL approach. I think it is well described. However, if the other effects (biochemical parameters, exophthalmos, etc) are considered as a whole, there may be sufficient data for a BMD analysis. Further, given the limited data on alopecia, if this was chosen as the sole criterion for analysis, then it might be invalid to exclude other studies that seem to have been left out for very inconsequential reasons such as Wei, and Manzo.

*A5. Please comment on the selection of the uncertainty factors applied to the POD for the derivation of the RfD. For instance, are they scientifically justified and transparently and objectively described in the document? If changes to the selected uncertainty factors are proposed, please identify and provide a rationale(s).*

While the use of uncertainty factors and the actual factors selected is entirely standard practice the review feels compelled to remind the EPA that if a greater uncertainty was utilized, the analysis would be prohibited. The maximal allowable value of 3000 is suggestive of a very limited data set that is in fact, poorly and incompletely assessed. This reviewer feels compelled to question whether the value of 3000 was "forced" to allow for an analysis to move forward.

*A6. Please comment specifically on the database uncertainty factor of 10 applied in the RfD derivation. Please comment on the use of the database uncertainty factor specifically for the lack of adequate developmental toxicity studies and a two-generation reproductive toxicity study, and additional uncertainty associated with the limited data available on neurotoxicity in light of the potential for neurotoxicity to represent a sensitive endpoint for thallium exposure. 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?*

The database is so weak that the uncertainty factor of 10 must be utilized. This choice is clearly described and appropriate.

*(B) Inhalation reference concentration (RfC) for thallium and compounds No data are available to derive the RfC for thallium and compounds. The only published studies involving inhalation exposure include a few case reports (Hirata et al., 1998; Ludolph et al., 1986) that suggest an association between occupational exposure and toxicity (including alopecia, gastrointestinal symptoms, and neuropathy), but the route or routes of exposure in these workplace setting could not be established.*

*B1. Has the rationale and justification for not deriving an RfC for thallium been transparently described in the document?*

This reviewer agrees. Data here are far too insufficient to attempt any further analyses.

*(C) Carcinogenicity of thallium and compounds*

*C1. 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 thallium and compounds. Please comment on the scientific justification for the cancer weight of the evidence characterization. A quantitative cancer assessment was not derived for thallium. Has the scientific justification for not deriving a quantitative cancer assessment been transparently and objectively described?*

Once again this review agrees that the data are far too insufficient to make any conclusions about the carcinogenicity of thallium. This justification is clear and objective.

**Deborah Rice, Ph.D.**
Maine Center for Disease Control
Augusta, ME 04333

Dr. Deborah Rice received her Ph.D. in toxicology from the University of Rochester in 1976. She has over 20 years of experience in toxicology and neurotoxicology. She is currently a Toxicologist for the Environmental and Occupational Health Program of the Maine Center for Disease Control and Prevention, Department of Health and Human Services. She is also an Adjunct Professor in the Department of Applied Medical Sciences, and Member of the Center for Integrative and Applied Toxicology at the University of Southern Maine. She presently serves as Associate Editor for the professional journals *Environmental Research* and *Neurotoxicology,* as well as serving on the Editorial Board for *Neurotoxicology and Teratology*. She was a member of the Organizing Committee for the annual Neurotoxicology Conference from 1999 to 2004. She has served on numerous expert panels throughout the years, and has published research articles in peer-reviewed publications, such as *Environmental Health Perspectives, The Journal of Neuroscience, Neurotoxicology and Neurobiology, Neurotoxicology and Teratology, and Toxicology and Applied Pharmacology* to name a few.

**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?*

The document is generally clearly written. However, EPA rejected studies that may provide useful information for reasons that are inadequate.

p. 18, lns 30-32. No rationale is provided for the assertion that because alopecia was not observed in the same location in all animals, this is not likely an effect of the chemical. There is no reason to assume that hair loss would preferentially occur at a particular location rather than being individualistic.

pp. 24-25. Zasukhina *et al* study. The reasons for rejection of this study are weak. Although it would be preferable to know the number of males and strain of rat, it is unclear how lack of this information invalidates the study. Are data from some strains of rat acceptable and others not? A small number of females per group would be a problem if the study were negative, as lack of statistical power would be a concern, but that is not the case here. Moreover, 16-18 females per group are really not a particularly small number. Most egregiously, the fact that doses are lower than estimated human intake is irrelevant to evaluation of the science. Such a statement is completely inappropriate in a document on hazard assessment.

p. 25. Wei study. What is a "non-standard" strain of mice, and how would that be relevant? Are there data indicating that this strain is more or less sensitive than other strains for the endpoints under study? If not, this statement is irrelevant. Similarly, how is the study invalidated by not knowing the age of the mice? It is reasonable to assume that they were young adults rather than aged animals. How does not knowing terminal body weight compromise interpretation of the effects on sperm motility, dead sperm, and dead fetuses from untreated females? It seems that although these data would be interesting, they are irrelevant. The most problematic aspect of the study is the failure to report water intake. However, on p. 21 of this document (ln. 25) EPA assumes a specific food intake in a study that did not report intake with no such squeamishness. EPA could assume typical water consumption for the Wei study. Even if it is somewhat inaccurate, by the time the UFs are added and the RfD is truncated to a reasonable number of significant digits, it is unlikely to make a difference.

It appears that EPA did not want to consider these studies, and disregarded them without sufficient justification. These studies should have been compared to the MRI study with regard to dose-response (see below). EPA seems to be inconsistent with regard to criteria for inclusion of studies for hazard assessment. For example, the recent PBDE assessment relied on an experimental design that included a single dose to mice pups from a European strain (in some cases) and a statistical analysis that is considered invalid. Statement in this document concerning non-standard strain and nonstandard design are puzzling, considering that EPA uses whatever studies it considers to be appropriate in its various assessments. A guideline subchronic study is not in any way necessarily the best basis for an RfD derivation.

p. 44 ff. Section 4.5.4. Most of these studies do not address mechanism of action. They are observations, or address mode of action at best.

p. 49, lns 15-19. Alopecia is the typical diagnostic criterion because other agents, including other metals, produce gastrointestinal symptoms and polyneuropathy. The point of this paragraph is unclear, but implies that alopecia is somehow a critical most important effect, which is not true. The paragraph should be deleted.

p. 49, last paragraph. See comments above.

p. 53, lns 21-26. See above comments.

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

I am not familiar with the thallium literature, so I don't know whether critical studies have not been considered.

Two studies that address the issue of neurotoxicity:

P. Kennedy and J.B. Cavanagh, *Acta Neuropathol* **39**, 81-8, 1977.

A.S. Windebank, *Exp Neurol* **94**, 203-12, 1986.

EPA seems to dismissing missing hair on these animals if it is considered to be from grooming behavior rather than a direct effect on hair follicles. This ignores the fact that such "over-grooming" in any species may be a sign of stress, pain, or other changes in the central or peripheral nervous systems. An animal grooming itself bald is not a benign effect. Studies that address this issue include the following:

J.M. Greer and M.R. Capecchi, *Neuron* **33**, 23-34.

A.V. Kalueff, A. Minasyan, T. Keisala, Z.H. Shah, and P. Tuohimaa. *Beh Proc* **71**, 8-15.

A.V. Kalueff and P. Tuohimaa, *Behav Brain Res* **160**, 1-10.

J.M. Welch *et al*., *Nature* **448**, 894-900.

*3. Please discuss research that you think would be likely to increase confidence in the database for future assessments of thallium and compounds.*

Detailed study of the potential for thallium to produce neurotoxicity is necessary. In particular, a developmental neurotoxicity study is essential. This is of particular concern given that other metals that produce neuropathy and GI symptoms in adults (e.g. lead and arsenic) produce developmental neurotoxicity, including IQ loss in children. Studies of general developmental effects and reproductive toxicity are required. The ability of thallium to produce endocrine effects requires exploration, particularly given the evidence for thyroid disruption. The effects observed on serum biochemistry parameters should also be explored. Data on the relative toxicity and pharmacokinetics of the various thallium salts would be useful.

*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?*

I agree with EPA that confidence in the RfD should be low. This is not based on the divergent view of EPA on this assessment and that of the contractor in 1988. It is rather based on the lack of studies addressing the known toxic effects of thallium, particularly neurotoxicity, developmental toxicity, and endocrine effects. In addition, a NOAEL was not identified in the MRI study if all relevant endpoints are considered, which is always problematic. Further, studies rejected by EPA suggest that effects may occur at lower doses than in the MRI study. Although including these endpoints and studies would add confidence in the assessment, the database on thallium is limited, with significant gaps.

**Chemical-Specific Charge Questions:**

*A1. The 90-day oral gavage study by Midwest Research Institute (MRI, 1988) was selected as the basis for the RfD. 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.*

The 90-day oral gavage study may be an acceptable basis for the RfD, although this is not clear without further analysis of other studies. It is important to remember that this study failed to take advantage of the known effects of thallium poisoning in designing the study. The 1988 contract should have included specific measures of somatosensory and motor function, as well as sophisticated pathology of the peripheral and central nervous systems, rather than a study more suitable for a chemical about which nothing was known. Given that the relevant endpoints were for the most part not assessed, it is important that EPA consider the surrogate measures of potential nervous system dysfunction, as outlined below.

The EPA "A Review of the Reference Dose and Reference Concentration Processes" (2002) recommends deriving reference values based on suitable endpoints and studies individually, including relevant UFs, rather than the "critical study-critical effect" approach. It is therefore recommended that RfDs be derived from the MRI study (numerous endpoints, see below), and the Wei and Zasukhina *et al*. studies, and the results compared. EPA will then have information on where the MRI study lies with respect to sensitivity, given that it is a screening study. This strategy was used by the NAS in its review of the toxicity of methyl mercury for EPA, for example, in which numerous endpoints from three studies were submitted to BMD analysis. Even if endpoints from the contract study are used to derive the RfD, analyzing the other studies that have dose-effect data will provide information regarding the range of doses at which effects were found, and thereby address the issue of how confident one can be about the RfD derivation.

*A2. Alopecia (hair loss) was selected as the most appropriate critical effect for the RfD. EPA characterizes alopecia as being an adverse effect. Please comment on whether the science and mode of action information supports alopecia as an adverse effect. EPA has stated: "Whether alopecia is itself an adverse effect merits consideration. In humans, alopecia is generally reversible upon cessation of thallium exposure. Alopecia, however, appears to be a part of a continuum of dermal changes observed following thallium exposure, as well as one of a spectrum of effects on target organs that include the nervous and gastrointestinal systems. For these reasons, alopecia supported by two cases of hair follicle atrophy is considered an adverse effect." Please comment on whether the selection of this critical effect has been scientifically justified. Is EPA's choice transparently and objectively described in the document? Please provide a detailed explanation. Please identify and provide the rationale for any other endpoints that should be used instead of alopecia to develop the RfD.*

It is appropriate to consider alopecia as an adverse effect. The fact that it is on a continuum of adverse effects is really not the point, however. Alopecia is itself an undesirable condition, and whether it is reversible following cessation of exposure is irrelevant. The RfD is designed to be a dose that may be ingested daily without appreciable risk of harm; therefore such a line of reasoning does not make sense. Alopecia represents an overt toxic effect, not a measure on a mechanistic pathway, for which one could be concerned about reversibility. Let's use the effect of lead on blood pressure as an illustration. Even a small increase in blood lead levels results in an increase in blood pressure, which is associated with an increase in myocardial infarction and death. EPA has monetized this relationship in its justification for keeping lead out of gasoline. The fact that blood pressure may be reversible if lead levels are lowered is irrelevant to the argument.

The lead example also serves to illustrate the issue of the effects of small changes in an outcome on the population level. The changes observed in clinical biochemistry in the MRI study may be important, for example, and are completely ignored. It is important to understand the difference between reference ranges for a population (or sample) and normal range for an individual, which is likely to be much narrower. It is also important to understand the difference between individual risk and population attributable risk. Individual risk increases as the value for a particular outcome gets farther outside of optimum range, whereas population attributable risk will be greatest at moderate ranges, because far more people will be within this range and therefore suffer the adverse consequence (of high blood pressure, high cholesterol, etc). EPA recognizes this very well; for example, regulating on the basis of a small shift in IQ or blood pressure. Yet here there is apparently no appreciation in this assessment of the fact that shifting a population or sample distribution may have important public health consequences.

Also see discussion below concerning inclusion of other endpoints.

*A3. At the high dose in the MRI (1988) study, two female rats exhibited moderate to severe alopecia that could not be attributed to self-barbering or normal cyclic hair growth. Histologic examination of skin samples from these high-dose females showed atrophy of hair follicles. EPA considered these findings to be adverse, and thus the high dose in this study (0.25 mg/kg-day thallium sulfate) to be the lowest-observed-adverse-effect level (LOAEL). The mid-dose group (0.05 mg/kg-day thallium sulfate) was identified as the no-observed-adverse-effect level (NOAEL). Is EPA's interpretation of the study findings scientifically justified? Has this interpretation of the findings been transparently and objectively described in the document?*

*As part of the evaluation of alopecia as a critical effect for the RfD, EPA performed a series of Fisher's Exact Tests to determine if the incidence of alopecia in any of the three dose groups was statistically significantly elevated above controls using all cases of alopecia reported by MRI (1988). Please comment on whether EPA chose the appropriate data set and the appropriate statistical test for this analysis.*

*The study investigators reached a different interpretation of the study findings than did EPA. The investigators considered alopecia to be attributable to the cyclic pattern of hair growth in rodents and, consequently, did not consider these findings to be biologically significant. The high dose (0.25 mg/kg-day thallium sulfate) was identified in MRI (1988) as the NOAEL. Is the study authors' conclusion that the high dose (0.25 mg/kg-day thallium sulfate) represents a NOAEL justified and supported by the study data?*

For alopecia, the middle- and high-dose groups were both significantly different from controls. EPA states that it considers alopecia to be an adverse effect. Yet it considers the middle dose to represent a NOAEL, with no explanation. It seems that EPA really considers hair follicle atrophy to be the adverse effect. Since this was only examined at the high dose, only the high dose could be considered to exhibit an adverse effect under this criterion. Thus the data from the other dose groups is ignored, and EPA might as well have considered studies with only one dose. Since atrophy was examined only at the high does, it should be considered a LOAEL and an UF of 10 applied. In other words, EPA's approach is illogical.

Of greater importance, however, is the fact that other endpoints collected in the study are completely ignored. First, the so-called "barbering" produced by dosed animals cutting and presumably eating their own fur is dismissed as irrelevant. This represents a change in behavior, and should be considered an adverse effect. It could be caused by pain, an increased arousal level, or stress, for example. This behavior increased in a dose-response manner. There are a number of other endpoints that also have a clear dose-response relationship. These include rough coat, piloerection, shedding, lacrimation, exophthalmos, and miosis (see graph). In addition, if the incidence of behavioral measures of increased arousal or pain (tense/aggressive, hyperactive, self-biting, and vocalization) are added, this also represents an orderly dose-response relationship. (The caveat is that some of these counts may be from the same animal, but EPA has the individual animal data and can collate this appropriately.) These ignored endpoints represent changes in arousal level or other nervous system function, pain or discomfort, changes in thyroid function, and/or other signs of poor health. All of these effects are increased at the lowest dose compared to the two control groups.

This begs the question: If these are not to be considered adverse effects, why were they quantified? These endpoints are included in the functional observation battery (FOB) developed by EPA to evaluate gross nervous system function and overall health. Further, EPA presumably requested of the contractor that these endpoints be assessed. Why, then, are they considered irrelevant in this assessment?

The failure of EPA to acknowledge these signs is particularly surprising given the effects of thallium poisoning in humans, which include peripheral neuropathy, pain in the extremities and other parts of the body, paresthesias and numbness of the extremities, symptoms of central nervous system

impairment, and impaired peripheral nervous system function as measured electrophysiologically. Many of the endpoints in the rat study are consistent with these effects in humans, and should not be dismissed as not representing adverse effects. In addition, some of these effects were observed in other animal studies, which adds to the confidence that these findings are meaningful. The lowest dose is clearly an effect level for a number of endpoints, and should be considered a LOAEL.

As an additional point, the Methods for the study does not state that animals were actually removed from their cages for the "neurotoxicological observations", so it is likely that they were not. This would preclude any meaningful observation of gait, arousal level, coordination, etc. In fact, the study protocol is missing a number of important details of methodology that would be required for publication of the study in a peer-reviewed journal. Performing a study under GLP does not have any correlation with whether it is a good or poor study, but is supposed to ensure that one can determine what was done. This study fails. In addition, a more modern study would include some measure of motor activity even in this screening study.

As for the use of Fisher's exact test, it is a standard procedure for pair-wise comparisons. It fails to take advantage of the fact that there are three dose groups. A more appropriate analysis would be a trend analysis.

*A4. The traditional NOAEL-LOAEL approach was used to define the point of departure (POD) for the RfD. A benchmark dose (BMD) analysis was considered but was not conducted because of the nature of the data set for alopecia. Please provide comments with regards to whether a NOAEL-LOAEL approach is the best approach for determining the POD. Has the approach been scientifically justified? Is it transparently and objectively described? Please identify and provide a rationale for any alternative approaches for the determination of the POD, and if such approaches are preferred to EPA's approach.*

EPA should try the BMD approach for each of the endpoints that exhibit a dose-response relationship in the MRI study, as well as the Wei and Zasukhina *et al*. studies. If results are satisfactory, EPA should use the BMD analysis to identify a POD. The POD may be the lowest BMDL, the one that EPA thinks is the most reliable or most orderly under analysis, or some kind of average. If the LOAEL/NOAEL approach from the MRI study is used, the lowest dose should be the POD as a LOAEL.

*A5. Please comment on the selection of the uncertainty factors applied to the POD for the derivation of the RfD. For instance, are they scientifically justified and transparently and objectively described in the document? If changes to the selected uncertainty factors are proposed, please identify and provide a rationale(s).*

The UFs for animal-to-human and intrahuman variability are appropriate. An additional UF to go from a LOAEL to a NOAEL is required if the BMD approach is not used. If the lowest dose is used as a LOAEL, it may be that an UF factor of 3 should be used to protect against the possibility of effects at

lower doses with longer exposure. The problem is that there are so many data gaps in the thallium literature that the total UF would be unacceptably high if the typical approach is used.

*A6. Please comment specifically on the database uncertainty factor of 10 applied in the RfD derivation. Please comment on the use of the database uncertainty factor specifically for the lack of adequate developmental toxicity studies and a two-generation reproductive toxicity study, and additional uncertainty associated with the limited data available on neurotoxicity in light of the potential for neurotoxicity to represent a sensitive endpoint for thallium exposure. 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?*

A database UF of 10 may not be necessary if all the endpoints in the MRI are considered, some of which may be indicative of neurotoxicity. The lack of a robust database on developmental toxicity is problematic, however, particularly since there is evidence in animals and humans of developmental effects. In addition, developmental neurotoxicity and endocrine studies have apparently not been performed. There are some data providing evidence for reproductive toxicity, but only for one sex. Again, the three studies should be used to generate sample RfDs from the endpoints that exhibit a dose-effect function. A database UF of 3 may be more appropriate under that scenario.

However the UFs are parsed, a total UF greater than 3000 should not be applied. A greater total UF suggests that there is not enough information to derive an RfD.

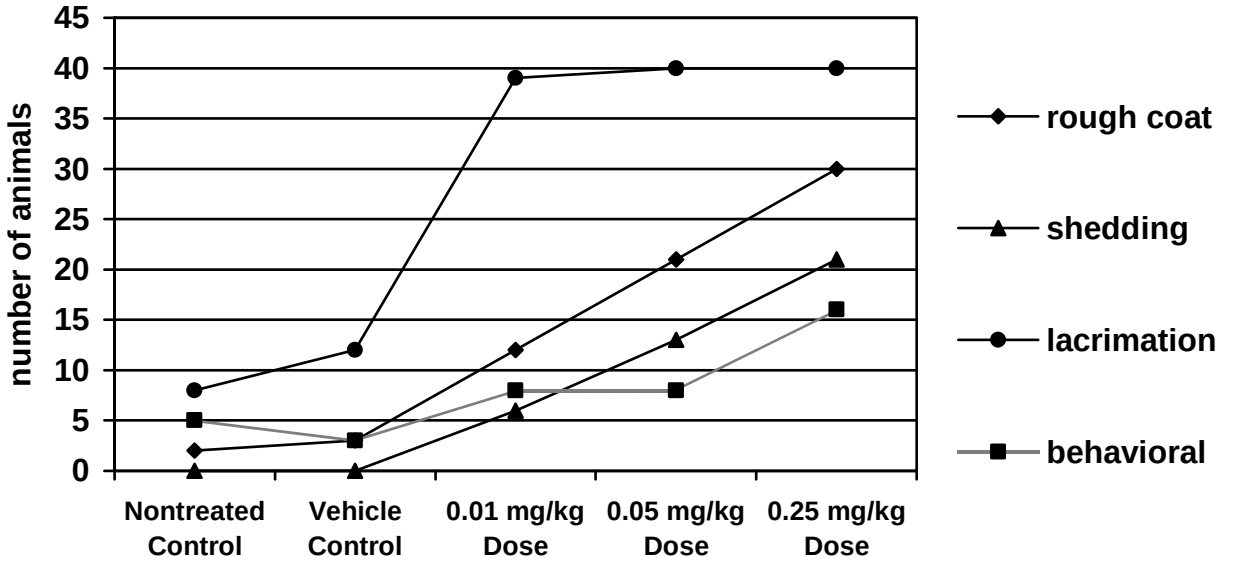
*B1. Has the rationale and justification for not deriving an RfC for thallium been transparently described in the document?*

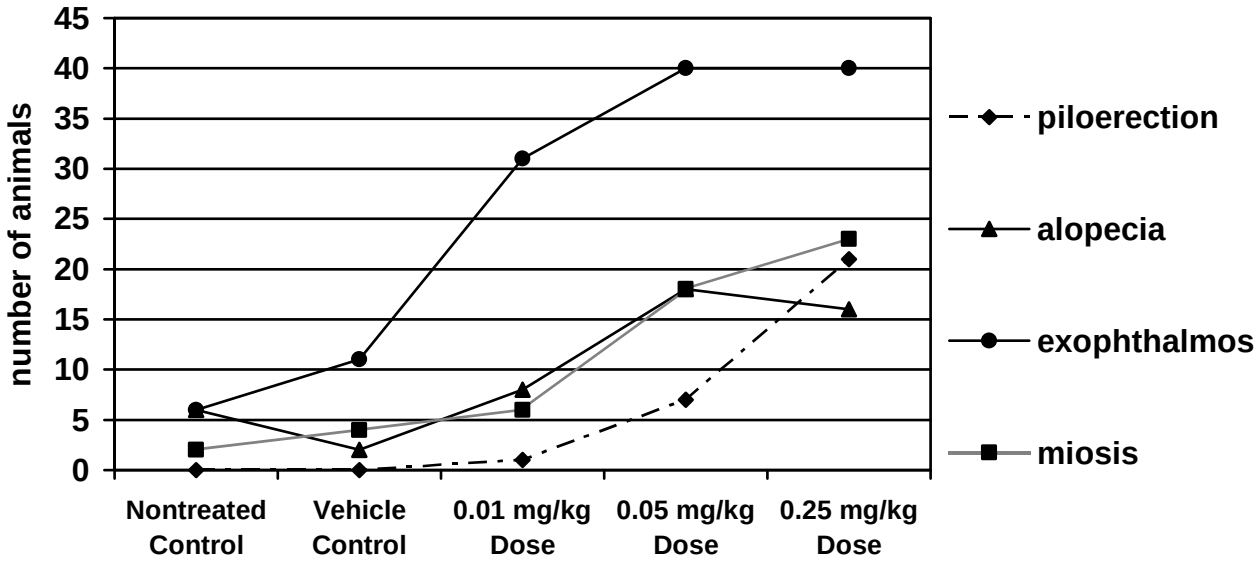
Given the relatively short half-life of thallium and lack of a PBPK model, it is reasonable that an RfC not be derived.

*C1. 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 thallium and compounds. Please comment on the scientific justification for the cancer weight of the evidence characterization. A quantitative cancer assessment was not derived for thallium. Has the scientific justification for not deriving a quantitative cancer assessment been transparently and objectively described?*

The rationale for not performing a quantitative risk assessment for cancer is appropriate.

**Incidence of specific effects out of a total of 40 animals**



# Appendix A

# List of Reviewers

# Peer Review Workshop for EPA's Draft Toxicological Review of Thallium and Compounds

Navy League Building
Arlington, VA
**May 19, 2008**

## Reviewer List

**Ronald Baynes**
Associate Professor
College of Veterinary Medicine
North Carolina State University
4700 Hillsborough Street
Raleigh, NC 27606
919-513-6261
Fax: 919-513-6358
Email: ronald_baynes@ncsu.edu

**George Cherian**
Professor Emeritus
Department of Pathology
University of Western Ontario
London, Ontario. N6A 5C1
Canada
519-661-2030
Fax: 519-661-3370
Email: mcherian@uwo.ca

**Lucio Costa (Chair)**
Professor
Department of Environmental and
Occupational Services
University of Washington
4225 Roosevelt #100
Seattle, WA 98105
206-543-2831
Fax: 206-685-4696
Email: lgcosta@u.washington.edu

**George Daston**
Research Fellow
Miami Valley Laboratories
The Procter & Gamble Company
Cincinnati, OH 45253
513-627-2886
Fax: 513-627-0323
Email: daston.gp@pg.com

**Robert Hoffman**
Associate Professor
Departments of Emergency Medicine and Medicine
New York University
455 First Avenue - Room 123
New York, NY 10016
212-447-8153
Email: bobhoff@pol.net

**Deborah Rice**
Toxicologist
Maine Center for Disease Control
State House Station 11
Augusta, ME 04333
207-287-3365
Fax: 207-287-3981
Email: deborah.rice@maine.gov

# Appendix B

# Observers

# Peer Review Workshop for EPA's Draft Toxicological Review of Thallium and Compounds

Navy League Building
Arlington, VA
**May 19, 2008**

## Observers

**Jan Connery (Facilitator)**
Vice President
ERG
110 Hartwell Avenue
Lexington, MA 02421
781-674-7322
Fax: 781-674-2851
Email: jan.connery@erg.com

**Karen Hammerstrom**
Deputy Director, IRIS Program
Office of Research and Development - IRIS
National Center for Environmental Assessment
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460
703-347-8641
Email: hammerstrom.karen@epa.gov

**Kathleen Newhouse**
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460
703-347-8641
Email: newhouse.kathleen@epa.gov

**Susan Rieth**
Toxicologist
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW (8601P)
Washington, DC 20460
703-347-8582
Email: rieth.susan@epa.gov

.

# Appendix C

# Meeting Agenda

# Peer Review Workshop for EPA's Draft Toxicological Review of Thallium and Compounds

Navy League Building
2300 Wilson Boulevard
Arlington, VA
**May 19, 2008**

## Agenda

8:00 a.m. **Registration**

8:30 a.m. **Welcome, Introductions, Meeting Purpose & Agenda**..... *Jan Connery, ERG*

8:40 a.m. **EPA Welcome Remarks**..... *EPA/NCEA Management Representative*

8:50 a.m. **Public Comment**..... *Jan Connery*

9:00 a.m. **Discussion Process and Overarching Comments**..... *Lucio Costa (Chair) & Panel*

9:15 a.m. **Oral RfD for Thallium** ..... *Lucio Costa & Panel*

A1) Does the document scientifically justify the choice of MRI (1988) as the principal study? Does it describe the study transparently and objectively? If a different study should be used, please provide the rationale.

A2) Do the science and mode-of-action information support alopecia as an adverse effect? Is selection of alopecia as the critical effect for the RfD scientifically justified and transparently and objectively described? If a different endpoint should be used, please provide the rationale.

10:15 a.m. BREAK

10:30 a.m. **Oral RfD for Thallium (cont.)**..... *Lucio Costa & Panel*

A3a) Is EPA's interpretation of the MRI (1988) high dose (0.25 mg/kg-day thallium sulfate) as the LOAEL and mid dose (0.05 mg/kg-day thallium sulfate) as the NOAEL scientifically justified? Does the document transparently and objectively describe this interpretation?

A3b) Did EPA choose the appropriate data set and statistical test when it used the Fisher's Exact Tests to determine the statistical significance of alopecia incidence?

A3c) Is the MRI (1988) authors' conclusion that the high dose (0.25 mg/kg-day thallium sulfate) represents a NOAEL justified and supported by the study data?

A4) Is EPA's NOAEL-LOAEL approach the best approach for determining the POD? Has the approach been scientifically justified and transparently and objectively described? Provide a rationale for any alternative approaches to determine the POD, and clarify whether such approaches are preferred to EPA's approach.

**Agenda (cont.)**

A5) Are the uncertainty factors applied to the POD for derivation of the RfD scientifically justified and transparently and objectively described? Please provide a rationale for any proposed changes to the selected uncertainty factors.
A6) Was selection of the database uncertainty factor of 10 for the RfD derivation scientifically justified? Is it transparently and objectively described?
Summary of Key Reviewer Comments on Oral RfD for Thallium

Noon LUNCH

1:00 p.m. **Inhalation RfC for Thallium and Compounds** ..... *Lucio Costa & Panel*
B1) Is the rationale and justification for not deriving an RfC for thallium transparently described?

1:15 p.m. **Carcinogenicity of Thallium and Compounds** ..... *Lucio Costa & Panel*
C1) Please comment on the scientific justification for the cancer WOE characterization that there is *inadequate information to assess the carcinogenic potential* of thallium and compounds. Does the document transparently and objectively describe the scientific justification for not deriving a quantitative cancer assessment?

1:30 p.m. **General Questions** ..... *Lucio Costa & Panel*
G1) Is the document logical, clear and concise? Does it accurately, clearly and objectively represent and synthesize the scientific evidence for noncancer and cancer hazard?
G2) What additional cancer and noncancer studies should EPA consider for thallium and compounds?
G3) What research would increase confidence in the database for future assessments of thallium and compounds?
G4) Sections 5 and 6: Does the document adequately discuss the key sources of uncertainty? Does it transparently and objectively describe the choices and assumptions made and the impact of uncertainty on the assessment?

2:45 p.m. BREAK

3:00 p.m. **General Questions (cont.)** ..... *Lucio Costa & Panel*
G4) Sections 5 and 6: Does the document adequately discuss the key sources of uncertainty? Does it transparently and objectively describe the choices and assumptions made and the impact of uncertainty on the assessment?

3:20 p.m. **Reviewer Final Comments**..... *Lucio Costa & Panel*

4:25 p.m. **Closing Remarks** ..... *Jan Connery & EPA/NCEA*

4:30 p.m. ADJOURN