

**Comments of the ACC Hydrocarbon Solvents Panel**  
**Attachment I**

**Justification for considering the “pain sensitivity” neurotoxicity endpoint as evidence of acute CNS effects with no persistent effects with continuous exposure**

The comments of the ACC Hydrocarbon Solvents Panel are organized according to EPA’s Charge Questions to the Science Advisory Board (SAB) for the draft IRIS Toxicological Review of Trimethylbenzenes (Draft Assessment). These comments focus on accuracy, objectivity and transparency of EPA’s analyses of primarily the neurotoxicity endpoints because they are considered to be the critical effect for risk assessment.

**GENERAL CHARGE QUESTIONS**

**1. NRC (2011) state that “all critical studies need to be thoroughly evaluated with standardized approaches that are clearly formulated” and that “strengthened, more integrative, and more transparent discussions of weight of evidence are needed.” NRC also indicated that the changes suggested would involve a multiyear process. Please comment on EPA’s success thus far in implementing these recommendations.**

ACC Comments:

EPA’s summary tables in the text and in Appendix B are significant improvements over the manner in which data was tabulated and presented 5 to 10 years ago. The presentation of information makes it easier to review EPA’s assessment of the studies. As will be discussed in greater detail in response to Charge Question C1, there were inaccuracies and serious omissions of key results in the summary tables in Appendix B, which translated into incorrect evaluation of the weight of evidence. In addition, the summaries do not capture important methodological limitations that affect study quality and interpretation of consistency of results. From this perspective, the guidance for evaluating the weight of evidence in the preamble has not been fully utilized. This discussion focuses on the section on neurotoxicity because the point of departure is based on pain sensitivity.

**4. EPA solicited public comments on the draft IRIS assessment of trimethylbenzenes and has revised the assessment to respond to the scientific issues raised in the comments. A summary of the public comments and EPA’s responses are provided in Appendix F of the Supplemental Information to the Toxicological Review of Trimethylbenzenes. Has EPA adequately addressed the scientific issues?**

ACC Comments:

EPA has not adequately addressed earlier comments from ACC that the C9 toxicity studies should be included in the trimethylbenzene (TMB) review. Although EPA argues they have a risk management need to regulate TMBs independent of other aromatic hydrocarbons (F-2), it does not follow that all mixed C9 isomer (i.e., approximately 55% TMB, 28% ethyl toluene isomers) toxicity studies should be dismissed as irrelevant to the scientific weight of evidence for TMB. In fact, EPA included discussion of findings from animal and human neurotoxicity studies on other substances, including toluene, xylene and white spirit, which are less closely related to the TMB isomers than the commercial mixed C9 isomeric substances (pp 1-1, 1-2, 1-23; 2-7). Compared to white spirit (which contains 15-25% aromatic

constituents), the C9 aromatic naphtha used in the neurotoxicity study is more relevant to TMB risk assessment because it contains ~ 87% aromatic hydrocarbons with nine carbon atoms, primarily ethyl toluene (ET; 28%) and TMB (53%) (Douglas et al. 1993). The reason this is an important point is that the C9 aromatic naphtha neurotoxicity study, in particular, contributes data on latency to paw lick in the hot plate test, which is directly relevant to evaluating the reversibility of the critical endpoint that was selected as the basis for the point of departure.

EPA's rationale for excluding the mixed C9 isomer studies stems from the EPA's observation that "multiple peer-reviewed studies have been published that demonstrate that individual TMB isomers do elicit clearly adverse toxicological effects", whereas the C9 fraction studies generally "failed to observe clear measures of toxicity in the systems investigated" (F-3). In other words, EPA made an *a priori* assumption that the findings from primarily one laboratory are incontrovertible, and excludes the mixed C9 isomers studies because of largely negative results (F-3) in the neurotoxicity studies evaluating the potential for persistent effects. This is not an objective approach to evaluating the weight of evidence; nor is this correct. For the neurotoxicity studies, it is more likely that the differences in findings are related to different behavioral tests, duration of exposures, time of test relative to last exposure, whether time of testing was balanced across dose groups, and the extent to which experimental bias and environmental factors affecting behavioral testing were controlled. Additional rationale supporting the relevance of mixed C9 isomer studies to the TMB risk assessment will be discussed below in comments to Charge Question B1 below.

### Chemical-Specific Charge Questions

#### B. Literature Search Strategy/Study Selection

**1. Please comment on the whether the literature search approach, screening, evaluation, and selection of studies for inclusion in the assessment are clearly described and supported. Please identify any additional peer-reviewed studies from the primary literature that should be considered in the assessment of noncancer and cancer health effects of 1,2,3-TMB, 1,2,4-TMB, and 1,3,5-TMB.**

#### ACC Comments:

**A detailed justification for the inclusion of the studies on complex C9 aromatic substances is provided in Attachment II.**

The EPA clearly describes the literature search approach, screening, and selection of studies for inclusion in the assessment, **but does not provide any rationale for excluding studies on complex C9 solvent mixtures, such as C9 aromatic fraction containing TMBs in the main body of the report**. A rationale can be found buried in Appendix F "Resolution of Public comments" of the Supplemental Material, but this is not transparent unless included in the Executive Summary and the Literature Search Strategy Section. EPA's decision to exclude C9 studies lacks consistency and objectivity because EPA includes papers that do not meet EPA's criteria for relevant test chemicals, such as Hissink et al. (2007); Battig et al. (1956, 1958); Chen et al. (1999); Lammers et al. (2007) (rats and humans exposed to white spirit); Lee et al. (2005); Norseth et al. (1991); and Sulkowski et al. (2002).

As discussed in our response to General Charge Question 4, a risk management need to regulate TMB alone does not imply that relevant scientific literature on similar group of chemicals or solvent mixtures should be ignored, particularly when these groups of chemicals are present together in the environment. The scientific basis for EPA's decisions will be strengthened if the mixed C9 isomer studies currently

summarized in Appendix E of the supplementary material are evaluated similarly to those included in Appendix B and integrated into the appropriate sections of the main report. Just as EPA indicates that the rats in the Lammers et al. (2007) study were exposed to 1,2,4-TMB as a constituent of white spirit (Table B-34; p. B-123), EPA can indicate that the rats in the Douglas et al. (1993); McKee et al. 1990 and Schreiner et al. (1989) studies were exposed to 1,2,4-TMB as a constituent of C9 Aromatic Naphtha. This will improve the objectivity and transparency of EPA’s selection of studies for inclusion in the assessment and include all the relevant scientific evidence.

EPA suggests that the negative results of the C9 mixture toxicity studies could be due to interactive effects between the constituents of the C9 mixture and that biological systems could alter the ADME of TMB (F-13 line 4). This hypothesis has not been tested directly, but acute toxicity and PK data suggest that this is not likely to be the case. In a 3-day acute exposure study by McKee et al. (2010), a complex C9 aromatic solvent produced similar, but perhaps more profound effects than did 1,2,4-TMB in tests of operant visual discrimination, functional observation battery and motor activity. This study clearly documented the timing of behavioral observations (functional observational battery and motor activity) relative to end of exposure and counterbalanced the dose groups across time of testing and testing device for the operant 2-choice visual discrimination performance test. Based on a pilot PK study in humans, Järnberg et al. (1998) concluded that exposure to 1,2,4-TMB in white spirit appear to increase the concentration of 1,2,4-TMB in blood and its metabolites in urine compared to exposure to 1,2,4-TMB alone. Since it is known that aromatic constituents induce their own metabolism, this apparent difference was likely due to increased metabolism of 1,2,4-TMB with co-exposure to other C9 aromatics in white spirits, including other TMB isomers. Apart from this, the disposition of 1,2,4-TMB did not differ whether exposure was to 1,2,4-TMB alone or in white spirit (Tables 2 and 3 in Järnberg et al.,1998). Thus, from an acute toxicity and pharmacokinetic perspective, these data support EPA’s earlier assumption that “assessing the toxicity of the C9 mixture as a complete entity should provide a reasonable upper bound” for the toxicity of TMB [isomers] in the C9 mixture (EPA, 1985).

In addition, EPA’s PBPK modeling validation and optimization suggest that there may be only modest impacts of other constituents of white spirit, a hydrocarbon solvent (approximately 80% aliphatic and 20% aromatic) containing more constituents other than TMB compared to complex C9 aromatic solvents. The PBPK model EPA uses for the Draft Assessment is based on a model developed by Hissink et al. (2007) following single day of exposure of rats and humans to white spirit containing only 7.8% TMB (after spiking) and having an overall aromatic content of 25.6%. Relatively modest changes in  $V_{max}C$  and  $K_m$  values were needed to optimize fit of this PBPK model to produce acceptable simulation of venous blood 1,2,4-TMB for repeated exposures to 1,2,4 TMB by inhalation. In fact, EPA reported that a  $V_{max}C$  value of 3.39, which is 3% different from Hissink’s original value, improved the model fit for humans exposed to TMB only, and was not significantly different from EPA’s selected value of 4.17.

Table 1. Chemical specific parameters used for each step to optimize the white spirit PBPK model for exposures to 1,2,4-TMB only.

	Original Hissink et al. 2007 parameters (8 hr exposure to white spirit)	Step 1. Optimization to Hissink et al. 2007 (8 hr exposure to white spirit)	Step 2. Optimization to Swiercz et al. 2003 (6 hr/day, 5 days/4 weeks to TMB only)	Step 3. Optimization to human data exposed to TMB only
EPA reference	Table B-5, B-6; pp. B-21-26	p. B-34	Table B-7; p B-36	p. B-43

V <sub>max</sub> C (% difference from 3.5) mg/hr/kg <sup>0.70</sup>	3.5	3.08 (-12%)	<b>4.17 (20%)</b>	<b>4.17 (20%)</b>  3.39 (-3%) (improved model fit, but not significantly from 4.17)
Km (% difference from 0.25) mg/L	0.25	0.050 (80%)	<b>0.322 (20%)</b>	<b>0.322 (20%)</b>

Note: % difference is in comparison with original Hissink et al. 2007 PBPK model parameters based on TMB in white spirit; final EPA model parameters are in bold.

In summary, EPA’s decision to exclude mixed C9 isomer studies lacks accuracy, objectivity and transparency when considering that EPA appropriately included animal and human studies for solvent mixtures and other solvents in the weight of evidence *and also as the fundamental basis for PBPK modeling for TMB alone*. In addition, direct comparison of the acute central nervous system (CNS) effects of exposure to individual C9 isomers with that from exposure to a complex C9 substance show no substantial differences. Thus, EPA’s rationale for excluding the C9 mixture studies is not supported by the available data. **As stated earlier, a detailed justification for the inclusion of the TSCA Section 4(a) studies conducted on complex C9 aromatic substances is provided in Attachment II.**

### C. Hazard Identification

#### *Synthesis of Evidence*

**1. A synthesis of the evidence for trimethylbenzene toxicity is provided in Chapter 1, Hazard Identification. Please comment on whether the available data have been clearly and appropriately synthesized for each toxicological effect. Please comment on whether the weight of evidence for hazard identification has been clearly described and scientifically supported.**

#### ACC Comments:

EPA’s summary tables in the text and in Appendix B are clearly presented. However, the summaries do not capture important methodological limitations that affect study quality and interpretation of results. Table 2 of these comments highlights experimental design issues and summarizes available historical control data. In addition, EPA’s reporting of statistical significance is inaccurate in several instances. Table 3 summarizes the results of statistical analysis including important results of the main ANOVA analyses that EPA did not include in any of the Supplementary or Main report summary tables. Table 3 also reports the post-hoc comparisons between treatment level and control because in some cases EPA mistakenly reported within group statistical significance as between group statistical significance. Due to these issues, the weight of evidence for the neurotoxicity section has not been accurately described, and the guidance for evaluating the weight of evidence in the preamble has not been followed.

Of special concern is that all of the studies by Korsak and colleagues did not indicate if time of testing was balanced across exposure level and devices or if the subjective measures were conducted by the same or different observers without knowledge of exposure level. In addition, the original studies did not clearly report all the statistical analyses conducted for each of the variables. In some cases statistical

comparisons were made across trials within a treatment group (i.e. L3 vs. L1) and inferences were made regarding significance across a treatment group (treated L3 vs control L3). This is confusing and may have contributed to errors in EPA's summary table.

Regarding the Korsak and Rydinski (1996) study, ACC appreciates EPA's need to select a sensitive endpoint that is amenable to BMD analysis and agrees that both acute and long-term effects of repeated exposures to TMB are important. However, ACC does not agree with EPA's conclusions regarding consistency of the acute pain sensitivity finding with effects reported 50-51 days after repeated exposure for 4 weeks. This is important because it impacts the selection of uncertainty factors. (See in particular response "2" below in which we show that latency to paw lick in the hot plate tests, a measure of pain sensitivity, is different from the results of the hot plate-foot shock tests, which are a measure of active avoidance.)

1. **The EPA's tables do not accurately report the statistical results or the total number of comparisons.** Comments below are focused on EPA's tables for studies conducted by Korsak, Gralewicz, Wiaderna and colleagues on pain sensitivity and conditioned analgesia. Based on the types of errors EPA made for these endpoints, EPA should check summaries of statistical results for other endpoints involving repeated trials. The statistical analyses conducted by Korsak and colleagues included 2-way ANOVA (4 dose groups and 3 trials as factors; reported as repeated measures in some but not all papers); 1-way ANOVAs for each trial across dose groups (to investigate significant interactions) and ratios of trials; and post-hoc comparisons of the direct measures (L1, L2, L3) and ratios (L2/L1, L3/L1). The post-hoc comparisons were conducted between (e.g. TMB vs. control) and within (e.g. L3 vs L2 for each TMB level) treatment groups. At a minimum, the Supplementary Appendix B tables should be revised to include significant and non-significant results of the 2-way and 1-way ANOVA analyses (considered to be the main analyses by the authors) and post-hoc comparisons, and include explanation of whether significant findings are comparisons between or within groups and if they are for direct measures or ratios. For this purpose, we summarize the results of ANOVA analyses focused on comparisons between control and treated groups (Table 3).
  - 1.1. **Table 1-1 for 1,2,4-TMB Gralewicz et al. (1997) (p. 1-10) incorrectly reports that statistical significance was achieved for latency after trial 3 (L3).** Only the ratio of L3/L1 was statistically significant when comparing high (250 ppm) and mid (100 ppm) doses with controls. The direct measure L3 was not reported by Gralewicz et al. (1997) to be statistically significant in any for the exposed groups compared to controls.
  - 1.2. **Table 1-1 for 1,2,4-TMB Gralewicz and Wiaderna (2001) (p. 1-10) is accurate, but more discussion is needed in the report text.** EPA's tabulation of data as percent of concurrent control can be misleading to the casual reader if the data are not discussed rigorously in the report. For example, the post-shock trial 3 (L3) is reported as a 191% difference for Gralewicz and Wiaderna (2001) but the pre-shock value (L1) was also 206% of control. This indicates a lack of an effect of the shock to produce an analgesic effect but an overall increase in latency across all trials. However, these results conflict with those from Gralewicz et al. (1997). This lack of consistency is discussed and illustrated in greater detail below (Figure 1).
  - 1.3. **Table 1-1 for 1,2,3-TMB Wiaderna et al. (1998) (p. 1-12) incorrectly reports statistical significance at the mid-dose compared to controls for response at 51 days post-exposure (24 hr. after foot shock).** The report is inaccurate because there was no statistically significant difference between the mid-dose and control group in the direct measurement L3. The significance was in reference to L1 (i.e. L3 vs. L1) within each treatment group and not in comparison to the concurrent control L3. This error is discussed in greater detail below.

- 1.4. **EPA's table B-42 Figure 4 from Wiaderna et al. (1998) (1,2,3-TMB: hot plate immediately before and after foot shock) is not transparent because EPA does not report important negative results and does not provide sufficient explanation of statistically significant results.** EPA did not include the footnote for the upper panel that “\*p<0.05 compared with L1 in the same group” (note that the original authors correctly indicated p<0.5). EPA also did not include the results that were reported in this study as not statistically significant, namely: “Post hoc comparisons revealed no differences between groups within successive trials”. This led to the error in EPA's summary Table 1-1 (p. 1-12) of the main report, incorrectly indicating significance at the mid dose compared to control.

**Table 2. Description of and Control Data for Pain Sensitivity and Conditioned Analgesia Studies**

		Age or weight of animal at start of study	Duration of exposure	Test time interval since last exposure	Shock from test prior to hot plate test	Hot Plate Shock immediately after Trial 1	Balance time of testing across dose level	Latency trial 1 control mean/error	Latency trial 2 Control mean/error	Latency trial 3 Control mean/error
<b>Pain (Thermal) Sensitivity</b>										
<b>1,2,4-TMB</b>	<b>Korsak, 1996</b>	250-300g	3 mo	Immediate	Rotarod 2mA	No shock	No	15.4 / (s.d.)=5.8 EPA assumed s.d.	n.a.	n.a.
<b>1,2,3-TMB</b>	<b>Korsak, 1996</b>	250-300g	3 mo	Immediate	Rotarod 2mA	No shock	No	9.7 / (s.d.)=2.1 EPA assumed s.d.	n.a.	n.a.
<b>Solvents</b>	<b>Korsak, 1994</b>	330 g	3 mo	Immediate	Rotarod 2mA	No shock	No	12.2/ s.d.=3.1	n.a.	n.a.
<b>C-9 aromatic naphtha mixture (55% TMB)</b>	<b>Douglas, 1993</b>	300 g	3 mo	2 days	None	No shock	Yes, random	Wk 0, 5, 9, 13, respectively: 8 / s.d.=2.7 12.2 /s.d.=4.8 10.2/ s.d.=3.8 10.9/s.d. = 4.2	n.a.	n.a.
								Wk 0 lab historical control 9.5-12.2		
<b>THERMAL SENSITIVITY (Trial 1) and CONDITIONED ANALGESIA WITH SHOCK (Trial 2, 3)</b>										
<b>1,2,4-TMB</b>	<b>Gralewicz 1997</b>	5 mo	4 wk	50-51d	Passive Avoidance 100 ms 4 mA: 1 Hz 10s	100 ms 2 mA: 0.5 Hz for 2 min	No	23 / (sem)=3	44 / (sem)=4	25 / (sem)=4
<b>1,2,4-TMB</b>	<b>Gralewicz 2001</b>	5 mo	4 wk	50-51 d	Passive Avoidance 100 ms 2 mA: 1 Hz for 10s	100 ms, 2 mA: 0.5 Hz for 2 min	No	9 / sem=1	41 / sem=6	11 / sem=2
<b>1,2,3-TMB</b>	<b>Wiaderna 1998</b>	5 mo	4 wk	50-51 d	Passive Avoidance 100 ms 2 mA: 1 Hz for 10s	100 ms, 2 mA: 0.5 Hz for 2 min	No	16 / (sem)=4	34 (sem)=4	18 (sem)=5
<b>1,3,5-TMB</b>	<b>Wiaderna 2002</b>	5 mo	4 wk	50-51 d	Passive Avoidance 100 ms 4 mA: 1 Hz 10s	100 ms, 4 mA: 0.5 Hz for 2 min	No	22 / (sem)=2	43 (sem)=3	22 (sem)=3

**Table 3. Statistical results of Pain sensitivity (Hot Plate) and Conditioned Analgesia Tests (Hot plate paired with shock)**

		2-way ANOVA (4-dose;3trial; sometimes reported as repeated measures ANOVA)			1-way ANOVA (4 dose)			Group comparisons with control			Group comparisons with control	
TMB isomer	First author, year	Dose	Dose x trial	trial	L1	L2	L3	L1	L2	L3	L2/L1	L3/L1
1,2,4 acute effect	Korsak 1996	n.a.	n.a.	n.a.	-	n.a.	n.a.	*100↑ *250↑	n.a.	n.a.	n.a.	n.a.
1,2,4	Gralewicz 1997	Θ <sup>a</sup> [L1&L2]	Θ <sup>a</sup> [L1&L2]	* <sup>a</sup> [L1&L2]	(Θ)	(Θ) [ANOVA L2/L1: Θ]	(Θ) [ANOVA L3/L1: *]	(Θ)	(Θ)	(Θ)	(Θ)	*100↑ *250↑
1,2,4 (100 ppm <sup>b</sup> )	Gralewicz 2001	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	Θ	Θ	*100↑	-	-(Θ)
1,2,3 acute effect	Korsak 1996	n.a.	n.a.	n.a.	-	n.a.	n.a.	*25↑ *250↑	n.a.	n.a.	n.a.	n.a.
1,2,3	Wiaderna 1998	Θ	(*) <sup>c</sup>	*	Θ	Θ	Θ	Θ	Θ	Θ	(Θ)	(Θ)
1,2,3 (100 ppm <sup>b</sup> )	Gralewicz 2001	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	Θ	Θ	*100↑	-	-(Θ)
1,3,5 (100 ppm <sup>b</sup> )	Gralewicz 2001	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	Θ	Θ	Θ	-	-
1,3,5	Wiaderna 2002	Θ	*	*	Θ	Θ	*	Θ	Θ	*100↑	Θ	Θ
C9 (55% TMB)	Douglas 1993	n.a.	n.a.	n.a.	-	n.a.	n.a.	Θ	n.a.	n.a.	n.a.	n.a.

**Notes:** Korsak and colleagues reported numerous comparisons within a group (i.e. Trial 1 vs Trial 3 within each exposure level) and these are not reported in this table. EPA's tables sometimes mistakenly reported within group comparisons as between group comparisons (control vs. treated). For example, Wiaderna et al. 1998 did not report any group comparisons with control to be significant, analysis was always between trials within a dose group. This table only reports post-hoc comparisons conducted between treatment group and controls. This table also reports results of ANOVA that were reported to be conducted in at least one of the publications.

<sup>a</sup> 2-way ANOVA reported for trials 1 and 2. It is likely that analyses for all 3 trials were also conducted but not reported.

<sup>b</sup> Only 100 ppm level tested for each test chemical. This study tested multiple chemicals, but only one exposure level.

<sup>c</sup> For 1,2,3-TMB, Wiaderna et al. 1998 reported that the dose x trial interaction was not statistically significant, but based on *p* value it is likely to be significant

**L1, L2, L3** = latency for trials 1, 2, 3, respectively, for the conditioned analgesia test with shock immediately after trial 1;

– “authors did not report results of an analysis, and it is not clear if this analysis was conducted

“**n.a.**” statistical analysis is not applicable to the experimental design because there was only one trial

**Θ** not statistically significant

**(Θ)** It is likely analysis was conducted based on methods section and/or post-hoc tests reported, but statistically insignificant results were not reported

\* indicates the ANOVA was significant for the specified factor or interaction

The columns for post-hoc comparisons of treated with control lists the dose levels that were statistically significant and indicates the direction of change from control

- 1.5 EPA's Table B-43 Figure 2 for Wiaderna et al. (2002) (1,3,5-TMB : hot plate immediately before and after foot shock) did not report the results of the 2-way ANOVA and large number of mostly non-statistically significant comparisons.** Of the 27 comparisons (9 between group for each trial + 12 within group comparisons of trials + 6 between group for proportions L2/L1 and L3/L1), the only significant result was an increase in L3 when comparing the mid-dose with control and low-dose. This should not be considered an effect on pain sensitivity because there were no significant increases in L1 at any dose level compared to controls. In the results section the authors state, "in none of the groups did the reaction latencies in trial 3 differ significantly from those determined in trial 1" and "no significant differences were detected between the groups in the values of the proportions L2/L1 and L3/L1". This indicates that there were no effects on the conditioned analgesia test, and the biological significance of the increase in L3 at the mid-dose level is uncertain.
- 2. EPA's assessment of pain sensitivity conflates two different behaviors involving hot plate into a single "pain sensitivity phenotype". This incorrectly gives the impression of persistency and consistency of finding.**

The test of "Pain Sensitivity" (pp. 1-2 to 1-4) should be divided into two sections – one on "pain sensitivity" that discusses the hot plate test without shock, and the other on "conditioned analgesia" for the hot plate combined with shock paradigm.

The pain sensitivity test evaluates the response to noxious stimuli (hot temperature) as an unconditioned stimulus. In the hot plate shock paradigm, only the first of three trials is identical to the pain sensitivity test because the animals have not been shocked. Trials 2 and 3 are models of "conditioned analgesia" in which an unconditioned (shock) and/or possibly conditioned (hot plate environment) aversive stimuli are used to induce analgesia before or during exposure to a noxious stimulus (hot temperature) (Butler and Finn, 2009; Miguez et al. 2014). There are different hypotheses regarding whether an increase in latency during Trial 2 or 3 is due to fear, stress, or improved memory (for trial 3). Although these hypotheses are speculative, none of these behaviors would be considered relevant to the pain sensitivity test that is measured without shock. These behaviors should be discussed separately in different sections entitled "pain sensitivity" and "conditioned analgesia".

Indeed, one of the EPA's external peer-reviewers for the 1995 Koch Industries study report on the 90-day oral toxicity study of 1,3,5-TMB<sup>1</sup> highlighted this exact same inconsistency in the final report submitted to the EPA. In discussing his findings regarding the Korsak "pain sensitivity" studies, the peer reviewer suggested that the Korsak studies on "pain sensitivity" with the addition of footshock "*may not be a valid or equivalent paradigm usage for classical conditioning*". It is not clear why this reviewer's comments were ignored by the EPA.

There were a large number of statistical comparisons conducted by Korsak, Gralewicz, Wiaderna and colleagues. Based on a review of all of these papers, the standard approach by this laboratory is to conduct 2-way ANOVAs, large number (up to 27) different post-hoc comparisons within and between groups and additional 1-way ANOVAs. The authors focused on reporting the significant findings. Table 3 selects only the group comparisons between treatment and control groups because this is the standard of comparison for risk assessment purposes.

---

<sup>1</sup> Peer Review Report – External Peer Review of the 1995 Koch Industries Study Report. 90-Day Oral Gavage Toxicity Study of 1,3,5-Trimethylbenzene in Rats with a Recovery Group. Page 19.

3. **Gralewicz and colleagues did not replicate their own results for the effects of 100 ppm 1,2,4-TMB on pain sensitivity.**

Figure 1 below illustrates the lack of concordance in the results from two different experiments on the effect of 100 ppm 1,2,4-TMB on pain sensitivity and conditioned analgesia. The results of the statistical analyses are reported in Table 3. The Gralewicz and Wiaderna (2001) study (left panel of Figure 1) shows that 1,2,4-TMB increases pain sensitivity (L1; not statistically significant from control) but has no effect on conditioned analgesia because the ratio of L3 to L1 are relatively the same for control and chemical. The Gralewicz (1997) study (right panel) shows a different result. There is no effect of 1,2,4-TMB on pain sensitivity, but there is an effect of the shock (i.e. effect on conditioned analgesia) based on differences between L3 and L1. The Gralewicz (2001) study measured the effect of one single concentration (50 ppm) of xylene, 1,2,4-TMB, 1,2,3-TMB and 1,3,5-TMB. In this study, the results for 1,2,4-TMB indicated there were increases in L1 that altered the “baseline” behavior just prior to the shock. This change in pre-shock “baseline” (L1) was not observed in the 3 other studies evaluating the effects of multiple doses on these same TMB isomers. There were no significant effects of 1,2,4-TMB on the ratio L3/L1, which therefore indicates there is no effect of TMB on conditioned analgesia. Thus, there is conflicting evidence of the effect on conditioned analgesia at 100 ppm. This, together with limitations in study design (discussed below) reduces the utility of this endpoint for risk assessment.

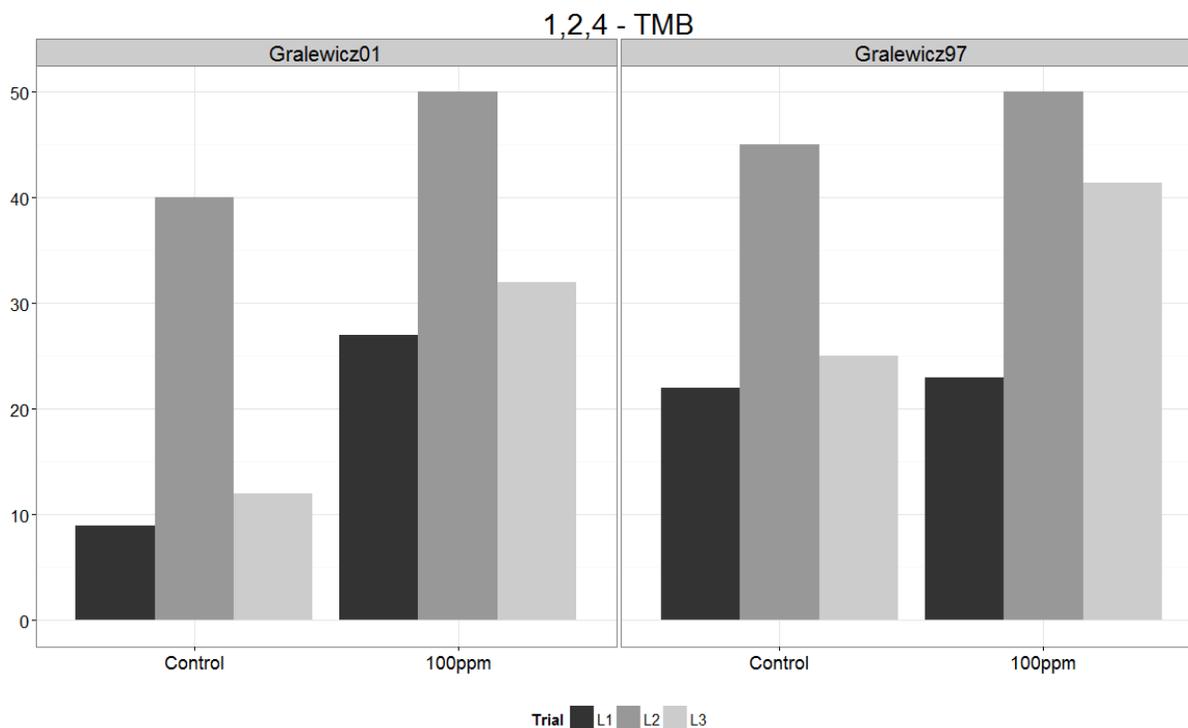


Figure 1: Results from two studies that show inconsistent results for the effect of 100 ppm 1,2,4-TMB. Data were estimated from the graphs of the original papers. Statistical results are summarized in Table 3 of these comments. Although there is an apparent increase in L3 in both studies, the pattern of effects are inconsistent with each other. The left panel shows no effect on conditioned analgesia, the right panel shows an effect on conditioned analgesia. The left panel shows a numerical increase in pain sensitivity (L1), the right panel does not.

**4. Interpretation of the biological significance of effects on L3 or L3/L1 is made more difficult by (a) the confounding effects of the passive avoidance test (animals associate leaving a platform with shock) that was completed just 2 days prior to the hot plate test and (b) the absence of a control “sham” group for the shock .**

The objectivity of this discussion can be improved by including discussion of limitations in interpreting changes in L3 or L3/L1 and overall lack of replication of findings (Table 3 of our comments). For example, an increase in L3 could reflect improved memory of the conditioned (hot plate environment) aversive stimuli that was associated with receiving a shock for two minutes 24 hours earlier. It could also be due, in part, to the fact that the rats were just tested two days earlier on a passive avoidance paradigm in which the rats were required to suppress their normal tendency to step down off of a platform. Following trial 3 of the passive avoidance test, the animals receive shocks for 10s. This prior testing with passive avoidance confounds the interpretation of the hot plate-shock test that uses the same aversive stimuli for 2 minutes (see Table 2). In addition, Wiaderna et al. (1998) observed that in the 1,2,4-TMB study “licking the hind-paw, was usually preceded by attempts to get out of the plastic enclosure. The more persistent were the attempts, the longer was the paw-lick latency.” This would suggest that the increased latency may not necessarily be an adverse effect. This reduces the level of concern for changes in L3 or L3/L1 parameters 50 days after four weeks of exposure to TMB, and increases confidence that the BMDL for the reversible Korsak et al. paw-lick finding will be protective of equivocal effects.

**5. EPA does not follow the guidance in the preamble to evaluate the quality of experimental studies including control of other variables that could influence the occurrence of effect, assessment of study quality characteristics and examination of historical control data from the same laboratory.** In general, the studies by Korsak and colleagues had several weaknesses in study design which were not described or considered in the weight of evidence:

**5.1. Subjective measures were not conducted blind to treatment level**

**5.2. The time of testing was not randomized (or balanced) across dose group (Table 2)**

Counterbalancing all the relevant factors within and across test sessions is a basic requirement of EPA’s neurotoxicity test guideline and expectation for quality behavioral studies. If the different dose groups were tested in sequential order (i.e. all control animals, then all low dose, etc.), statistically significant differences between treated and control group could be due to factors other than exposure level. The interpretation is confounded, and no statistical method could disentangle group and time effects (Maurissen, 2010).

**5.3. Historical control data from the laboratory indicates wide range of “normal” latency on hind limb paw lick (Table 2).** EPA should discuss the wide range of “normal” latency and discuss observations by Wiaderna et al. (1998) that animals that react to pain stimuli by trying to escape have longer latencies.

**5.3.1.**L1 values for 4-week exposure studies conducted in 5-month old animals and tested 50 days later range from **9 – 22** (Table 2 of our comments).

**5.3.2.**L1 values for 3-month studies conducted in rats weighing 250-300 g range from **9.7 -15.4**. Hence the experimental value of  $11.8 \pm 3.8$  for 25 ppm 1,2,3-TMB, although statistically different from the concurrent control value of  $9.7 \pm 2.1$  (Korsak and Ryzdzyński, 1996) falls within the range of control values and should not be considered as a treatment-related effect. These data support a NOAEL of 100 ppm for 1,2,3-TMB and 1,2,4-TMB because the L1 values were within this range of control values (Korsak and Ryzdzyński, 1996).

**5.3.3.**One of the reasons why latency to hindlimb paw lick (trial 1) can be variable is that the rats may have different strategies for escaping the aversive heat stimuli which is not captured by measuring hindlimb paw lick.

**6. EPA's discussion of behavior treats conjecture as scientific evidence for incorrect conclusions regarding behavioral measures. The most significant examples are in the Hazard Assessment sections on "Motor Function and/or anxiety" and "Pain Sensitivity". This severely weakens the scientific credibility of the TMB review.**

- 6.1. **The motor activity tests should not be interpreted as effects on anxiety.** The tests for motor function were not designed or validated to measure "anxiety". Thus changes in motor behavior cannot be interpreted as an effect on "anxiety". All references to "anxiety" should be removed from this section, including the title of this section.

**The Douglas et al. (1993) neurotoxicity study of C9 aromatic naphtha has relevant data for TMB risk assessment, including hot plate latency,** automated startle response, hind foot splay, grip strength and an especially strong evaluation of relevant tissues for neuropathology. The neuropathology included evaluation of perfusion fixed peripheral and central nerve tissue with H&E stains and luxol fast blue stain for myelin degeneration, cross and longitudinal sectioning and nerve teasing. The automated 30-minute motor activity test measured a large number of variables, but the data are too variable.

- 6.2. Our comments to General Charge Question 4 discuss in detail why EPA's decision to exclude this study from the TMB assessment is not scientifically sound or consistent with EPA's inclusion of studies involving white spirit.

- 6.3. **The Douglas et al. neurotoxicity study provides additional evidence that the increased latency to hind paw lick (without shock) is reversed within two days of approximately 4, 8 and 13 weeks of inhalation exposures to C9 aromatic naphtha mixture containing 55% TMB.**

***Summary and Evaluation***

- 1. Does EPA's hazard assessment of noncancer human health effects of trimethylbenzenes clearly integrate the available scientific evidence (i.e., human, experimental animal, and mechanistic evidence) to support the conclusions that trimethylbenzenes pose potential hazards to the nervous system, respiratory system, the developing fetus, and the circulatory system (i.e., blood)?**

ACC Comments:

The EPA's summary and evaluation is clearly written. There were no studies designed to measure effects of TMB isomers on anxiety, and reference to this should be omitted from line 9 of page 1-52. By excluding the C9 aromatic naphtha neurotoxicity study, EPA fails to integrate all the available relevant scientific evidence (see ACC comments to Charge Question B.1 for detailed discussion). This is important because the C9 aromatic naphtha neurotoxicity study indicates that there are no effects on hot plate hind limb paw lick when measured two days following 4, 8 or 12 weeks of exposure, which is consistent with the results of Korsak and Ryzdzyński (1996). In addition, there are no effects on a very thorough neuropathology evaluation following 3 months of exposure. **This data should be added to the weight of evidence because it reduces concern that the sensitive critical effect will increase in severity with increased duration of exposure.**

EPA's characterization of the TMB effect on the hot plate foot shock results give the impression that these results are the same type of effect as the hot plate results without the foot shock. EPA describes the hot plate-foot shock results as a measure of pain sensitivity following environmental challenge. This

would be akin to characterizing active avoidance as primarily a measure of activity following an environmental challenge. In addition, the potential confounding effect of the passive avoidance test just prior to the hot plate-foot shock test decreases the utility of the hot plate-footshock test for risk assessment.

**2. Does EPA's hazard assessment of the carcinogenicity of trimethylbenzenes clearly integrate the available scientific evidence to support the conclusions that under EPA's *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005), there is "inadequate information to assess the carcinogenic potential" of trimethylbenzenes?**

ACC Comments:

The charge question may be misleading. The EPA's *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005) require that all relevant information be reviewed. Although there are no cancer studies of individual TMB isomers, there are repeated dose (90 days, 12 months) of C9 aromatic naphtha in which there is no evidence of pre-neoplastic lesions (Clark et al., 1989). Additionally, the C9 aromatic naphtha was evaluated in a battery of *in vitro* and *in vivo* tests and found to be non-genotoxic (Schreiner et al., 1990). The EPA reviewed this information and concluded that no additional carcinogenesis testing was necessary (Meranda, 1988). Thus the statement that there is "inadequate information to assess the carcinogenic potential" of trimethylbenzenes does not convey the information that the relevant data were reviewed by the EPA and determined to be sufficient for hazard characterization purposes.

**D. Toxicokinetics and Pharmacokinetic Modeling**

**1. Please comment on whether the selected PBPK model (Hissink et al., 2007) with EPA's modifications adequately describe the toxicokinetics of 1,2,4-TMB (Appendix B). Was the PBPK modeling appropriately utilized and clearly described? Are the model assumptions and parameters scientifically supported and clearly described? Are the uncertainties in the model structure adequately characterized and discussed?**

**2. The internal dose metric selected for use in the derivation of the RfC and RfD for 1,2,4-TMB was the steady-state weekly average venous blood concentration (mg/L) of 1,2,4-TMB for rats exposed for 6 h/day, 5 days/week. Please comment on whether the selection of this dose metric is scientifically supported and clearly described. If a different dose metric is recommended for deriving the RfC, please identify this metric and provide scientific support for this choice. Are the uncertainties in the selected dose metric adequately characterized and discussed?**

ACC Comments:

The EPA's description of model assumptions and parameters are clearly described. We note that *when EPA optimized the PBPK model based on white spirit to fit data from TMB exposures only very modest changes in chemical-specific parameters were required (see detailed response to Charge Question B4). From a pharmacokinetic perspective, this would suggest the other constituents of C9 (other than TMB) do not have a major influence on the pharmacokinetics of TMB based on this indirect evidence.*

**E. Inhalation Reference Concentration (RfC) for 1,2,4-TMB**

**1. A 90-day inhalation toxicity study of 1,2,4-TMB in male rats (Korsak and Rydzyński, 1996) was selected as the basis for the derivation of the RfC. Please comment on whether the selection of this study is scientifically supported and clearly described. If a different study is recommended as the basis for the RfC, please identify this study and provide scientific support for this choice.**

**2. Decreased pain sensitivity (measured as an increased latency to paw lick response after a hotplate test) in male Wistar rats was concluded by EPA to be an adverse effect on the nervous system and was selected as the critical effect for the derivation of the RfC. Please comment on whether the selection and characterization of this critical effect is scientifically supported and clearly described. If a different endpoint(s) is recommended as the critical effect(s) for deriving the RfC, please identify this effect and provide scientific support for this choice.**

**3. In order to characterize the observed dose-response relationship comprehensively, benchmark dose (BMD) modeling was used in conjunction with dosimetric adjustments for calculating the human equivalent concentration (HEC) from a rat and human PBPK model (Hissink et al., 2007) to identify the point of departure (POD) for derivation of the RfC. Please comment on whether this approach is scientifically supported for the available data, and clearly described. a. Has the modeling been appropriately conducted and clearly described, based on EPA's *Benchmark Dose Technical Guidance* (U.S. EPA, 2012)? b. Has the choice of the benchmark response (BMR) for use in deriving the POD (i.e., a BMR equal to 1 standard deviation change in the control mean for the latency to paw lick response) been supported and clearly described?**

ACC Comments:

The use of the Korsak and Rydzyński (1996) study is a sensitive endpoint that is amenable to BMD analysis that will adequately be protective of the other behavioral effects that have been reported for TMBs by primarily one laboratory. The BMR equal to 1 standard deviation change in the control mean is appropriate given the wide variability in control values across studies.

However, results from a subchronic toxicity study do not necessarily indicate that the observed effects are persistent or increase in severity with exposure. For example, the hot plate latency to paw lick responses in the Korsak and Rydzyński (1996) studies are essentially the same whether the rats were exposed once for 4 hours or continuously for 3 months. In addition, the response observed after the 3 month exposure period is seen only when the rats are tested immediately after the last exposure. When tested 2 weeks after the last exposure, no statistically significant effects relative to control are reported by the authors. This response clearly indicates an acute reversible CNS effect and should not be confused with persistent effects based on difference in duration of exposure alone.

**4. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the POD for the derivation of the RfC for 1,2,4-TMB. Are the UFs appropriate based on the recommendations described in Section 4.4.5 of *A Review of the Reference Dose and Reference Concentration Processes* (U.S. EPA, 2002), and clearly described? If changes to the selected UFs are proposed, please identify and provide scientific support for the proposed changes.**

ACC Comments:

The additional subchronic to chronic UF of 3x is unnecessary because there is strong weight of evidence that there are no effects of repeated exposure on pain sensitivity as measured by latency to paw lick on a hot plate. Using EPA's guidance on weight of evidence, there are five dose response studies supporting the lack of a long lasting effect on paw-lick latency two weeks after a 4-week exposure (Korsak and Rydzynski, 1996; Gralewicz et al. 1997; Gralewicz and Wiaderna 2001; Wiaderna et al. 1998, 2002). In addition, a neurotoxicity study on C9 aromatic naphtha (55% TMB) reported no effect on pain sensitivity to hot plate following 4, 8 and 12 weeks of exposure (animals were tested after a two-day holiday to C9). This C9 mixture neurotoxicity study also did not find any treatment related histopathology findings based on a thorough evaluation that exceeded EPA test requirements by including longitudinal of peripheral nerves and spinal cord and teased nerve fibers. In the 1,2,4-TMB dose response study, which should carry greater weight than the studies with only one dose level of 1,2,4-TMB isomer, there were no long lasting effects on conditioned analgesia (or stress-induced analgesia) as reflected by lack of consistent statistically significant effects on the proportion L3/L1. Taken together, these data support a reduction of the subchronic to chronic UF from 3 to 1.

In addition, the EPA included a UF of 3x to account for database insufficiency. In support of this, the EPA cites the lack of a multigenerational reproductive/developmental toxicity study as a weakness in the database. However, this weakness only exists because the EPA has chosen to ignore the existing data on the complex C9 aromatic substance. As summarized in Appendix E of the supplement to the EPA draft assessment, a 3-generation reproductive toxicity study in mice, 2 developmental toxicity studies in mice and rats and one developmental neurotoxicity study in rats, in which the complex C9 aromatic substance was tested, is available. On this basis, the inclusion of the UF for database insufficiency is not justified.

**References**

- Bättig, K; Grandjean, E; Rossi, L; Rickenbacher, J. (1958). Toxicologische untersuchungen uber trimethylbenzol. *Archiv fuer Gewerbepathologie und Gewerbehygiene* 16: 555-566.
- Bättig, K; Grandjean, E; Turrian, V. (1956). [Health damage after continuous exposure to trimethyl benzene in a painting workshop]. *Soz Praventivmed* 1: 389-403.  
<http://dx.doi.org/10.1007/BF02031676>
- Butler RK, Finn DP. (2009). Stress-induced analgesia. *Prog Neurobiol* 88(3): 184-202.
- Chen, R; Dick, F; Seaton, A. (1999). Health effects of solvent exposure among dockyard painters: Mortality and neuropsychological symptoms. *Occup Environ Med* 56: 383-387.  
<http://dx.doi.org/10.1136/oem.56.6.383>
- Douglas, JF; Mckee, RH; Cagen, SZ; Schmitt, SL; Beatty, PW; Swanson, MS; Schreiner, CA; Ulrich, CE; Cockrell, BY. (1993). A neurotoxicity assessment of high flash aromatic naphtha. *Toxicol Ind Health* 9: 1047-1058.
- Gralewicz, S; Wiaderna, D. (2001). Behavioral effects following subacute inhalation exposure to m-xylene or trimethylbenzene in the rat: A comparative study. *Neurotoxicology* 22: 79-89.  
[http://dx.doi.org/10.1016/S0161-813X\(00\)00003-6](http://dx.doi.org/10.1016/S0161-813X(00)00003-6)

Gralewicz, S; Wiaderna, D; Tomas, T; Rydzyński, K. (1997). Behavioral changes following 4-week inhalation exposure to pseudocumene (1,2,4-trimethylbenzene) in the rat. *Neurotoxicol Teratol* 19: 327-333. [http://dx.doi.org/10.1016/S0892-0362\(97\)00001-9](http://dx.doi.org/10.1016/S0892-0362(97)00001-9)

Hissink, AM; Krüse, J; Kulig, BM; Verwei, M; Muijser, H; Salmon, F; Leenheers, LH; Owen, DE; Lammers, JH; Freidig, AP; McKee, RH. (2007). Model studies for evaluating the neurobehavioral effects of complex hydrocarbon solvents III. PBPK modeling of white spirit constituents as a tool for integrating animal and human test data. *Neurotoxicology* 28: 751-760. <http://dx.doi.org/10.1016/j.neuro.2007.03.005>

Järnberg, J; Johanson, G; Löf, A; Stahlbom, B. (1998). Toxicokinetics of 1,2,4-trimethylbenzene in humans exposed to vapours of white spirit: Comparison with exposure to 1,2,4-trimethylbenzene alone. *Arch Toxicol* 72: 483-491. <http://dx.doi.org/10.1007/s002040050532>

Korsak, Z; Rydzyński, K. (1996). Neurotoxic effects of acute and subchronic inhalation exposure to trimethylbenzene isomers (pseudocumene, mesitylene, hemimellitene) in rats. *Int J Occup Med Environ Health* 9: 341-349.

Korsak, Z, Wisniewska-Knypl, J, Swiercz, R. (1994) Toxic effects of subchronic combines exposure to n-butyl alcohol and m-xylene in rats. *Int J Occup Med Environ Health* 7: 155-66.

Lammers, JH; Emmen, HH; Muijser, H; Hoogendijk, EM; McKee, RH; Owen, DE; Kulig, BM. (2007). Model studies for evaluating the neurobehavioral effects of complex hydrocarbon solvents II. Neurobehavioral effects of white spirit in rat and human. *Neurotoxicology* 28: 736-750. <http://dx.doi.org/10.1016/j.neuro.2007.03.003>

Lee, CR; Jeong, KS; Kim, Y; Yoo, CI; Lee, JH; Choi, YH. (2005). Neurobehavioral changes of shipyard painters exposed to mixed organic solvents. *Ind Health* 43: 320-326.

Maurissen J. (2010) Practical considerations on the design, execution and analysis of developmental neurotoxicity studies to be published in *Neurotoxicology and Teratology*. *Neurotoxicol Teratol* 32(2): 121-3.

Meranda, J. (1988). Letter to Dr. Robert Drew, American Petroleum Institute, May 3, 1988.

McKee, RH; Lammers, JH; Muijser, H; Owen, DE; Kulig, BM. (2010). Neurobehavioral effects of acute exposure to aromatic hydrocarbons. *Int J Toxicol* 29: 277-290. <http://dx.doi.org/10.1177/1091581810365089>

Mckee, RH; Wong, ZA; Schmitt, S; Beatty, P; Swanson, M; Schreiner, CA; Schardein, JL. (1990). The reproductive and developmental toxicity of High Flash Aromatic Naphtha. *Toxicol Ind Health* 6: 441-460.

Miguez G, Laborda MA, Miller RR. (2014) Classical conditioning and pain: conditioned analgesia and hyperalgesia. *Acta Psychol (Amst)* 145: 10-20.

Norseth, T; Waage, J; Dale, I. (1991). Acute effects and exposure to organic compounds in road maintenance workers exposed to asphalt. *Am J Ind Med* 20: 737-744. <http://dx.doi.org/10.1002/ajim.4700200604>

- NRC (National Research Council). (2011). Review of the Environmental Protection Agency's draft IRIS assessment of formaldehyde. Washington, DC: National Academies Press. <http://www.nap.edu/catalog/13142.html>
- Schreiner, CA; Edwards, DA; Mckee, RH; Swanson, M; Wong, ZA; Schmitt, S; Beatty, P. (1989). The mutagenic potential of high flash aromatic naphtha. *Cell Biol Toxicol* 5: 169-188.
- Sulkowski, WJ; Kowalska, S; Matyja, W; Guzek, W; Wesolowski, W; Szymczak, W; Kostrzewski, P. (2002). Effects of occupational exposure to a mixture of solvents on the inner ear: A field study. *Int J Occup Med Environ Health* 15: 247-256.
- Swiercz, R; Wiaderna, D; Wasowicz, W; Rydzyński, K. (2003). Pseudocumene in brain, liver, lung and blood of rats after single and repeated inhalation exposure. *Int J Occup Med Environ Health* 16: 61-66.
- U.S. EPA (U.S. Environmental Protection Agency). (2012). Benchmark dose technical guidance. (EPA/100/R-12/001). Washington, DC: Risk Assessment Forum. [http://www.epa.gov/raf/publications/pdfs/benchmark\\_dose\\_guidance.pdf](http://www.epa.gov/raf/publications/pdfs/benchmark_dose_guidance.pdf)
- U.S. EPA (U.S. Environmental Protection Agency). (2005). Guidelines for carcinogen risk assessment [EPA Report]. (EPA/630/P-03/001F). Washington, DC: Risk Assessment Forum. <http://www.epa.gov/cancerguidelines/>
- U.S. EPA (U.S. Environmental Protection Agency). (2002). A review of the reference dose and reference concentration processes [EPA Report]. (EPA/630/P-02/002F). Washington, DC: Risk Assessment Forum, U.S. Environmental Protection Agency. <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=51717>
- U.S. EPA (U.S. Environmental Protection Agency). (1985). Identification of specific chemical substance and mixture testing requirements; Ethyltoluenes, trimethylbenzenes, and the C9 aromatic hydrocarbon fraction. *Fed Reg* 50: 20662-20677.
- Wiaderna, D; Gralewicz, S; Tomas, T. (2002). Assessment of long-term neurotoxic effects of exposure to mesitylene (1,3,5-trimethylbenzene) based on the analysis of selected behavioral responses. *Int J Occup Med Environ Health* 15: 385-392.
- Wiaderna, D; Gralewicz, S; Tomas, T. (1998). Behavioral changes following a four-week inhalation exposure to hemimellitene (1,2,3-trimethylbenzene) in rats. *Int J Occup Med Environ Health* 11: 319-334.