Peer Review Meeting of EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Reviewer Post-Meeting Comments

September 2, 2011

Contents

Responses to Charge Questions

(A) Toxicokinetics and PBPK Modeling

A PBPK model developed by EPA based on models by Ward et al. (1997) and Fisher et al. (2000) was utilized in the Toxicological Review of Methanol. This model is described in Section 3 and a detailed description of the EPA model modifications, evaluation, and application are found in Appendix B. The PBPK model modified by EPA can estimate internal dose levels due to exogenous methanol exposure (i.e., doses above background). This modified methanol PBPK model was first applied to predict internal doses in experimental animals under bioassay conditions. Benchmark dose (BMD) modeling, using internal doses as exposure metrics, was then used to identify internal-dose points of departure (PODs) from the animal data. Finally the human PBPK model was used to identify human equivalent concentrations (HECs) or doses (HEDs) for each internal-dose POD.

Note: Background methanol levels have been subtracted by study authors from most of the mouse and rat pharmacokinetic data and those background levels are not reported. Since the goal is to predict risk above background, the EPA subtracted background levels from the pharmacokinetic data where it was otherwise included, to obtain a consistent total data-set for use in developing the PBPK models. The underlying assumption is that non-cancer risks from methanol exposure are due to increases in the levels of methanol or its metabolites above background.

A1. Please comment on the scientific soundness of the PBPK model used in this assessment.

Reviewer	Comments
Burbacher	The assessment includes detailed reviews of PBPK models for mouse, rat, and humans. The reviews are comprehensive but it is not clear to this reviewer what the strengths and weaknesses are for each model and why the nonhuman primate model was not included in the final model development. Some clarification of the process for evaluating the usefulness of each model for the assessment and why the nonhuman primate model was not included would be helpful.
Byczkowski	By definition, science compiles, completes and systematically organizes knowledge, providing as detailed description of the investigated phenomena as currently possible. In contrast, science-based human health assessments and toxicological reviews aim at supporting the U.S. EPA regulatory activities designed to protect public health. Thus, the models used in risk assessment serve only a relatively narrow goal of estimating the health protective exposure levels of xenobiotics which are most likely without adverse effect to humans. To fulfill this goal, the modelers often have to choose shortcuts, simplifications, surrogate dose metrices, and to make health-protective assumptions for PBPK models, which may be unacceptable for scientific dissertation, but if technically correct - according to the rule of parsimony, they can be applied as tools in approximation of dose and/or response to xenobiotics for EPA regulatory purpose. The PBPK model used in this assessment includes at least three major shortcuts: i) it is lacking a detailed quantitative description/distribution of endogenous levels of methanol (background); ii) lacking qualitative and quantitative description of metabolites; and iii)

	lacking quantitative description of lactational transfer of methanol and its concentration in the postnatal brain. While such omissions could be perceived as a deficiency in scientific description of the phamacokinetics of methanol, this PBPK model seems to be adequate for the risk assessment purpose, for which it was developed. Obviously, this and any other PBPK model is only as good, as good were the data used for its validation. As stated in the reviewed document on P. 3-24 (L# 5 - 7): "...it was determined that a modified Ward et al. (1997) model, with the addition of a lung compartment as described by Fisher et al. (2000), should be used for the purposes of this assessment..." The choice and combination of these two PBPK models seems to be optimal. Even though, the combined "hybrid" model has been significantly simplified, it is still adequate for use in dose-effect modeling and interspecies extrapolations. The ACSL codes of the model have been listed in the Appendix B, including *.CSL and *.CMD files (and even the runtime *.m files). The PBPK model seems to be appropriately constructed, using the principle of parsimony, and it is very well documented. While the use of "drinking tables" (P. B-49; L#39 to P. B-50; L#12) cannot be considered to be "mathematically elegant", it represents a pragmatic solution to the problem of modeling different patterns of drinking/feeding by different species.
Dorman	There are several weaknesses associated with the use of the model at this time including: (a) lack of external peer review through publication of the EPA PBPK model in the open scientific literature; (b) the model does not explicitly consider gestational/lactational exposure and compartments; (c) there is no clear rationale provided as to when the model should be applied (e.g., prior to the dosimetric extrapolations or apply the PBPK model to a POD determined through the BMD approach); and (d) the model does not include any a description of metabolism to formaldehyde or formate. These weaknesses do not prevent the EPA from using the model. The model structure developed by the US EPA is based upon published models that were then adapted by the US EPA. The model structure is sound (although the use a bladder compartment is atypical – the EPA should consider recoding the model to include akidney/renal compartment that considers excretion of methanol by the kidney). The documentation provided by EPA is strong; however, the appendix describing the model contains extraneous information including email communications between scientists that do not contribute to a clear understanding of the model structure. Page 3-45 (andelsewhere) includes a description of the two divergent models that were considered (Michaelis-Menten or not) which I found confusing. The EPA should more clearly describe their reasons for developing and using the models. Page 3-49 – the description of the chamber volume should be expanded. I assume the equipment in question is the caging. Is that correct? Is there any evidence that incomplete mixing occurred since this group did examine methanol concentrations in different chamber locations prior to study start? Page 3-50. The EPA has not clearly articulated why two different fractional absorption values were used based on the same data base (see pages 3-50 (60%) and 3-42 (86.5%)). Page 3-50 – why was the second trimester group considered the most representative? This statement needs justification.

McMartin	The PBPK model that has been developed for this assessment is generally sound and has been thoroughly explained with generally appropriate justifications. An ideal scientific model of methanol pharmacokinetics/pharmacodynamics would have incorporated estimates of formate accumulation using parameters that estimated metabolism of formaldehyde (to formate and "other") and of formate (to CO_2 and urine formate excretion), similar to what was described for the Bouchard model. Incorporation of such parameters would be needed to help explain changes in blood methanol levels (since over long exposures, the rate of methanol elimination is governed by its metabolism). This, of course, would have made for a pretty complicated model. Because the EPA PBPK model is designed primarily to predict methanol blood levels across species (i.e., for HEC and HED calculations) for risk assessment purposes, the model presented in the document does a sufficient job. An interesting aspect of the developed model is that it has employed various parameters in order to maximize the fit to existing methanol blood level data and some of these parameters don't make physiological or biochemical sense. At first glance, the use of a saturable term for stomach absorption of methanol seems peculiar since the absorption most likely is a passive diffusion, first order process. However, absorption of methanol could become saturated at very high concentrations simply because these high levels slow down gastric emptying (like ethanol is known to do), thus limiting absorption in the intestine. Hence, the need for the saturable term can be understood. It is also strange that a bladder component would be needed for humans and not for rodents. I realize that the authors were attempting to model urinary methanol excretion since such data are available for human studies (and not in the published rodent studies), but it would be interesting to see how the model were changed if the same bladder component were included in the rodent models. The treatment of metabolism of methanol in the model is also hard to understand physiologically. For example, it is not clear why two saturable metabolic pathways are needed for the Sprague-Dawley rat and only one for the F344 rat in the sense that similar enzyme systems presumably operate in the two strains. It is understandable why the human model only incorporates one saturable pathway, because, at the low blood levels involved in this risk assessment, humans metabolize methanol exclusively by hepatic alcohol dehydrogenase (ADH) activity. Another odd result in the models is the apparent value of the biological constants (esp the K_m) that are calculated for the models (Table 3-10 and 3-11). These biological "constants" are actually calculated by the model through the curve fitting, that is the constants are changed iteratively in order to fit blood or urine level data from various published studies. The oddities lie in the fact that the resulting K_ms for the human and the rat range from 6 to 65 mg/L, whereas most biochemical studies with the rat and human ADHs report K_ms with methanol as a substrate in the range of 160 to 640 mg/L *in vitro* (which are similar to the values reported for primates *in vivo* in Table 3-11). It is understandable that the "constants" are varied to fit the model, but the constants don't seem to reflect the true Michaelis values of the metabolic enzymes themselves. In both these situations, it appears that the mathematics of the model is driving the biology, rather than having the biological explanations first, with the math following behind.
Roberts	Construction of PBPK models is outside my expertise. I see no obvious flaws in the model, but cannot comment on a technical level regarding its scientific soundness.

Salmon	The model is based on extensive and well-documented knowledge of the metabolism and kinetics of methanol and related compounds, including investigations of the differences between experimental animal species and humans. It also builds on the experience gained from several previous investigations of methanol toxicokinetics, and successfully avoids some limitations and peculiarities of these earlier models in developing a generally applicable model framework. The Agency is to be particularly commended for developing a consistent model framework and sets of species-specific parameters which have been validated across several somewhat diverse data sets. This considerably increases the confidence which may be placed in its conclusions, and is a welcome improvement on some earlier PBPK modeling exercises which did not test their validity by attempting to fit multiple independent datasets. The ability to effectively model the limited, but relevant, toxicokinetic data in humans as well as in the two experimental animal species of interest increases confidence in the usefulness of the model for interspecies extrapolation. One issue which needs attention is the fitting of the PBPK model to kinetic data for Sprague-Dawley (SD) rats. This is of significance for analysis of the NEDO developmental study. The analysis presented in the draft report relies on the dataset which appears most complete (which is an appropriate choice), but this results in an unexpectedly low value for the uptake fraction by inhalation for this strain of rat. It appears from discussions at the meeting that additional data from NEDO are now available which not only provide an alternative basis for parameterizing the model for SD rats, but also result in an uptake fraction similar to that seen for Fischer rats and mice, which is intrinsically more plausible biologically. This needs to be checked out in detail, and its consequences for the predicted HECs in the NEDO (1987) rat study explored.
Sweeney	<u>Rat Models</u> The Sprague-Dawley (SD) rat PBPK model is inappropriately parameterized (or insufficiently validated) for the inhalation route. The SD rat model was calibrated based on the iv and oral data. These data have advantages due to the greater certainty in delivered dose as compared to an inhalation study. However, no comparisons of the model predictions to the available inhalation data were made. The LOAEL of the key SD rat study was 1000 ppm. The simulations in Figure B-13 show a predicted blood Cmax of 10.6 mg/L at 1000 ppm (page B-26). The data in Table 3-5 (page 3-5), indicate a post exposure blood concentration of 83 mg/L in SD rats after 8 hrs exposure (Perkins et al., 1995a). (In contrast, on page 3-36 line 24, it is stated that "there are no inhalation data for SD rats".) If that "post exposure" blood concentration was taken immediately at the end of exposure, this datum suggests that the SD rat model is off by a minimum of a factor of 8 at this exposure concentration (if steady state had been achieved). In addition, pages missing from the PDF of the NEDO (1987) report were provided by the Methanol Institute during the review, and indicate thatF1 female SD rats exposed to 1000 ppm methanol had blood concentrations of 99.48 mg/L at the end of exposure. Likewise, the F344 rat model was not tested against the NEDO (2008b) data, where the post exposure blood levels for 1000 ppm exposures again exceed the steady state blood levels depicted in the 1200 ppm model fits (e.g., Figure B-11, page B-22). The conclusions for the F344 rat are of lesser concern for this assessment, since the (current) key toxicology study was conducted using SD rats.

	It should be noted that the EPA model uses (for both F344 and SD rats) a fitted fractional inhalation availability (FRACIN) derived for F344 rats (20%) which is much lower than the corresponding values for mice (66.5%) and humans (86.6%). In my use of the model, if FRACIN for SD rats is changed from 0.2 (20%) to 0.6 (60%), the 1000 ppm blood prediction is in agreement with Perkins et al. (1995a), suggesting that the Horton et al. F344 data set is an outlier. When the 60% value is used in the simulation of the kinetics in the NEDO developmental study, the predicted daily AUCs increase by 6.3-, 13.4-, and 22.1-fold at 500, 1000, and 2000 ppm respectively. I did not take this exercise so far as to redo the BMD analysis with the new internal doses, but given that the March 2011 draft BMDL was in the vicinity of the NOAEL (500 ppm), it seems reasonable to expect that a revised BMDL (and candidate RfC) would also increase by ~6 fold. EPA does not provide a sensitivity analysis of the rat PBPK model, even though PBPK-derived dose metrics from a rat study provide the basis of the RfC and RfD. A sensitivity analysis of the rat blood methanol AUC under conditions approximating the BMDL would appropriately focus the evaluation of model reliability on key model parameters. <u>Human Model</u> I have concerns regarding the parameterization/validation of the human PBPK model and the lack of human model sensitivity analyses, factors which undermine the confidence in the model and its application. Useful human kinetic studies were apparently overlooked by EPA and their contractors (see Section D2, below). While these studies are limited in the number of subjects involved, they are potentially quite valuable in model parameterization because they do not involve the inhalation route. The inhalation studies can have some substantial dosimetric uncertainties due to the possible interspecies differences in fractional uptake and uncertainty with respect to breathing rates. The human iv data (Haffner et al., 1992) and oral data (Schmutte et al., 1988) are not subject to the same dosimetric uncertainty. I also do not understand why EPA did not report any results from human model sensitivity analyses. Ideally, EPA would have conducted sensitivity analysis on steady-state blood methanol concentrations at the HEC and HED or RfC and RfD or values in between (see below, section A.2.). These analyses would focus the model confidence assessment on the parameters that are the key determinants of the internal dose and would inform the choice of UFH. At a minimum, EPA should assess whether or not the model they used in the risk assessment can (adequately) simulate the additional human data identified herein and conduct and provide human model sensitivity analyses at the RfC and RfD. EPA should further consider reparameterizing the human methanol PBPK model. <u>Mouse Model</u> It seems odd that, for oral dosing, the mouse blood levels are reported to be insensitive to any parameter related to clearance (e.g., metabolism, blood flow to the liver) (pp B-16 and B-18). It is also not clear what type of oral dose is being simulated based on the text alone (appears to be gavage in model files on-line). The runtime files that should reproduce Figures B-2 and B-5 yield simulations that are slightly off.

	EPA does not provide files that fully recreate the sensitivity analyses--only those parameters demonstrated in Figures B-6, B-7, and B-8. These files do not accurately reproduce the figures in the document. I tested an additional parameter that does not appear on the figures (FRACIN), and found that the model output was, as I expected, very sensitive to this parameter. Thus the sensitivity analysis does not appear to have been comprehensive.

A2. Please comment on the scientific justification for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks.

Reviewer	Comments
Burbacher	The rationale for this approach is weak. To be consistent with longstanding assumptions used by EPA and others, the critical factor related to toxicity is the total level of methanol from all sources, or cumulative exposure. The rational for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks is not clearly stated and it gives the impression that cumulative exposures from different sources are not important. On page 3-28 it states that the modeling that includes background levels was estimated to have "minimal impact" on the dose extrapolations. It would seem that if this is the case, these "more complex" models would also be more rigorous and appropriate for use in this assessment.
Byczkowski	The assumption that adverse effects of methanol exposure appears at internal concentrations higher than its physiological background is correct, analogously to many other chemicals - essential at physiological concentrations but deleterious at high external doses. The physiological levels of one carbon groups are metabolically necessary for the organism and they do not produce adverse effects. Since the U.S EPA can regulate only the external exposures - not the normal endogenous concentrations of chemicals, apparently a reasonable decision was made to subtract the background level from data used in the quantification of pharmacokinetics of methanol. On the other hand, any exposure to external methanol, adds up to the existing background, increasing concentration of methanol in the target tissue. So, the upper bound on background concentrations of methanol in target tissue should be carefully evaluated and used consequently. The lack of determination of the upper statistical bound on normal physiological concentrations of methanol in relevant species, including humans, can be considered to be a major deficiency of the reviewed document. The justification for selecting a no-background model, over the PBPK model that does include background, has been provided in the document (Section 3.4.3.2.1 and P. 3-28; L# 4 to 7): "...*more complex PBPK modeling required to include background levels was estimated to have a minimal impact on dose extrapolations, the use of simpler methanol models that do not incorporate background levels is considered adequate for the purposes of this assessment*..." Codes for the background level, blocked in the final simulations, have been incorporated in the PBPK model, as documented on P. B-48; L#9 to 55 and P. B-49; L# 1 to 3; and as stated on P. 6-3; L# 21 to 24: "...*This assessment focuses on the determination of noncancer risk associated with exogenous methanol exposures that increase the body burden of methanol or its metabolites (e.g., formate, formaldehyde) above endogenous*

	background levels..." However, in the simulations whose results are listed in the Table B-5, a background level of 2 mg/L has been set to model human internal concentration from inhalation (P. B-92; L# 29) but not from the oral exposure (P. B-92; L# 55).
Dorman	The available pharmacokinetic datasets are quite variable – as correctly noted by the EPA in some cases investigators reported "corrected" blood methanol concentrations (where baseline concentrations were subtracted) while in other cases this information was provided. EPA states that the impact of including endogenous methanol was minimal – that is likely correct until one approaches low blood methanol concentrations seen with exposures approaching the proposed RfC. In this case these concentrations approach or exceed the contribution that results from the additional exogenous methanol exposure. Ideally, the PBPK model should be revised to include endogenous methanol production/levels.
McMartin	The model has used background subtraction of endogenous methanol from the animal studies to prepare the PBPK calculation of HECs. By and of itself, the subtraction of background levels in the application of the PBPK model for calculation of risk assessment numbers does not appear to substantially affect the numbers obtained. However, consideration of whether to subtract background levels is extremely important in how the results are applied – in fact the way that this risk assessment treats endogenous and exogenous methanol levels is highly questionable. The assessment assumes that the endogenous levels of methanol (and its metabolites) do not contribute to the formation of adverse effects, which presumably is true. Although this assumption is scientifically justified, it creates a major problem for risk assessment of substances like methanol that are found endogenously. Basically, as is shown by the resulting RfC and RfD that are determined in this document, exposures of humans from the levels of methanol at the RfC or the RfD produce no increase in blood methanol above the endogenous background. If endogenous levels of methanol do not contribute to adverse effects and an exposure does not produce an increase above background levels, how can that exposure lead to an adverse effect? The conundrum occurs because the PBPK model itself has built-in conservatism, the BMD calculation has built-in conservatism and then a 100-fold uncertainty is applied. All of these factors contribute to bring the "RfC/RfD exposure" down to the levels where there is essentially no exposure-induced increase in methanol levels above the endogenous, background level, which means there is essentially no risk.
Roberts	The case for subtraction is based upon a stated goal of determining noncancer risk associated with exposures that increase body burden of methanol or its metabolites above endogenous levels. Two assumptions are stated: "(1) endogenous levels do not contribute significantly to the adverse effects of methanol or its metabolites; and (2) the exclusion of endogenous levels does not significantly alter PBPK model predictions." (pg 3-27). With respect to the second assumption, modeling with and without incorporation of background levels was tested using rat pharmacokinetic data with the stated result that incorporation of background had minimal effect (<1%) on the point of departure (POD). Given the doses of methanol used in the rat studies, the contribution of endogenous methanol to total blood levels was no doubt very small, and it is not surprising that the POD changed little whether or not endogenous methanol was included. The more important matter is the first assumption. If the proposed RfC and

	RfD values for methanol were associated with blood methanol levels much higher than endogenous levels, the contribution of endogenous levels would not be an issue. That is not the case, however. The RfC and RfD correspond to blood methanol concentrations in humans squarely in the range of normal "background" levels. I'm not aware of any evidence that endogenous and exogenous methanol are distinct in their potential to produce noncancer effects, and in fact there is no attempt to present a scientific argument on that point in the Toxicological Review. So, under the circumstances, the first assumption is not met.
Salmon	Conceptually this is a reasonable thing to do, since the non-cancer health effect data similarly consider the difference in incidence between controls (with the background methanol level) and exposed (with background plus exposure-related levels). Where there might be a problem with this is if the background levels were comparable to or larger than the exposure-related level, or if there was evidence of substantial perturbation of the background methanol metabolism by the additional exposure-related component. These concerns do not appear to be applicable at the exposure levels of interest in analysis of the experimental animal data and human studies analyzed for the development of the RfC. (The question of whether the background concentrations are comparable to levels achieved at the RfC/RfD is a different, and not necessarily relevant, question.) In particular, alternative model analyses in which the background concentrations were explicitly included produced very similar results to those looking at additional levels only. Since both approaches provide reasonable (and essentially identical) fits to the actual incidence data it appears that there are no major interactions between background and exposure-related levels to complicate the analysis of the animal datasets. There are modest differences in HEC and HED predictions with background included or excluded. It is not apparent from the narrative either in the main text or the appendix describing the model why these differences arise. I did not find any analysis of whether these differences should be considered significant from a statistical standpoint: it may be that there are insufficient human data to evaluate this question. On balance the Agency's decision to exclude the background levels from the calculations seems reasonable, although some further explanation is desirable. Apparently models both with and without inclusion of background levels have been developed, so that specific differences in these approaches could be identified. Including the background levels in the models necessarily increases the model complexity and like any model enhancement may increase the uncertainty in the final result, especially when as in this case it may be difficult to design a test of its validity. Part of the problem is the considerable variability and uncertainty in human background values for methanol. Also there is significant uncertainty as to mechanism and the extent to which individual genetic variability or adaptation may affect the possible toxicological significance of those values.
Sweeney	The subtraction of background levels (or the alternative modeling with background included) does not appear to have made a significant difference on the PBPK modeling of the available kinetic studies or BMD analysis. Having said that, it is my opinion, that if the object is to understand risk, and risk is related to internal dose, what we should be doing is attempting to define an acceptable (human) internal dose, and then, from that, derive acceptable external exposures. The sticking point with the methanol assessment is the interpretation of background levels and their contribution to risk. A fundamental difficulty, then, with this noncancer

	assessment, is that the internal dose BMDL (internal dose point of departure, iPOD) is converted to an HEC/HEC, and the uncertainty factors are applied to the HED/HEC to derive the RfD/RfC. Then, when the RfC/RfD is simulated, it is then realized that the additional methanol body burden is indistinguishable from background. A more direct route to the comparison to background is to take the iPOD, convert it to a time-weighted average (TWA) concentration (daily AUC divided by 24 hrs) and divide by the UFs directly (90.86 hr x mg/L/24 hrs = 3.8 mg/L; 3.8 mg/100 = 0.038 mg/L). It is immediately obvious that a difference of 0.038 mg/L methanol against a background of ~1.8 mg/L (with standard deviations of about 0.7 to 1.2 mg/L) is miniscule and thus is a de minimis increase in population risk. If you accept the UFA of 3, UFD of 3, and the pharmacodynamic component of UFH as 3, (I do not—see B.4), this gives you a composite of UF of 30 prior to consideration of human TK variability. If the acceptable additional blood methanol concentration above background is thus defined, with 3.8 mg/L ÷ 30 = 0.13 mg/L, the RfC would be ~5 ppm (see page B-37). I would like to see a sensitivity analysis on the *sum* of background methanol (~1.8 mg/L) plus additional methanol (0.13 mg/L) at 5 ppm. I doubt that total blood methanol would be very sensitive to any anticipated biological variability or parameter uncertainty under these circumstances. Thus the TK component of UFH can reasonably be 1. Even with a composite UF of 10 (e.g., with UFD = 1), the iPOD/UFc = 0.38, which is less than half the reported standard deviation of background methanol in most of the studies noted by EPA.

A3. The PBPK modeling effort assumed similar methanol pharmacokinetics between pregnant and non-pregnant animals. Please comment on the adequacy of the dose-metric extrapolation based on a PBPK model for non-pregnant adults (i.e., no fetal compartment) for predicting risks associated with fetal/neonatal brain concentrations of methanol.

Reviewer	Comments
Burbacher	The review of data related to methanol pharmacokinetics between pregnant and non-pregnant animals provides sufficient evidence for the assumption of similarity. There are data in nonhuman primates that indicate little change during pregnancy.
Byczkowski	Due to its physicochemical properties, "*...methanol penetrates cellular membranes readily and distributes throughout total body water...*" (c.f. P. 3-11; L# 14). This was confirmed by the experimental data, which according to the statement on P. 3-10 L# 20 to 22: "*... as a whole suggested that the distribution of orally and i.v. administered methanol was similar in rats versus mice and in pregnant rodents versus NP rodents...*" Figure 3-3 (P. 3-11) provided further evidence that within the relevant range of concentrations, the ratio of methanol concentration in amniotic fluid to concentration of methanol in maternal blood is nearly linearly 1 to 1. Thus, the dose metric extrapolation, based on non-pregnant adults to predict internal concentrations of methanol in fetal/neonatal brain, seems to be adequate, as explained and substantiated in the document.

Dorman	The limited pharmacokinetic data suggests that this assumption is valid. However, there remains a question as to whether this model can be applied to neonatal rats where blood methanol concentrations are > 2-fold higher than those seen in dams under similar exposure conditions. This issue is important since the critical study used by EPA to derive an RfC involved combined gestational and lactational (inhalational) exposure of neonates. The use of an adult-based PBPK model could under predict potentially 'toxic' blood methanol concentrations. Indeed, the RfC estimate may more closely approximate that obtained using a more standard approach that doesn't rely on a PBPK model.
McMartin	Existing literature strongly supports the assumption of similar pharmacokinetics between pregnant and nonpregnant animals (the Pollack and Brower as well as the Burbacher studies). Furthermore, studies have also indicated that fetal levels of the two major methanol metabolizing systems (ADH and catalase) are very low compared to adult levels, indicating that fetal tissues do not substantially impact the pharmacokinetics of methanol. As such, the use of a PBPK model based on non-pregnant adult data for predicting risks related to fetal concentrations of methanol can be substantiated for a dose-metric extrapolation. From a biological perspective, this makes sense – methanol is distributed evenly among tissues related to water content and so the levels in the fetal tissues should be similar if not identical as levels in the maternal tissues. Also, elimination of methanol controls its pharmacokinetics in most circumstances and the main driver for methanol elimination in the pregnant animal/human is the metabolic elimination by the maternal liver. Hence, the fetal levels of methanol are controlled mostly by the maternal liver, so the PBPK model is justified in utilizing nonpregnant data.
Roberts	Studies presented in the Review indicate that the pharmacokinetics of methanol are similar in pregnant and non-pregnant rodents, providing a scientific basis for modeling methanol concentrations in pregnant animals based upon data from non-pregnant adults. Given the data currently available, I understand the rationale for omitting a fetal compartment in the PBPK model. However, I think that for PBPK modeling to be effective, a fetal compartment will ultimately be needed. Studies such as Sweeting et al. (2011) suggest that maternal blood methanol concentrations alone are insufficient to explain developmental toxicity from methanol, even within the same species. PBPK modeling is most useful when the proximate form of the toxicant and mode of action are known, which is unfortunately not the case with developmental effects of methanol.
Salmon	The report describes various practical investigations and studies with PBPK models which indicate that, due to the free and rapid distribution of methanol throughout most maternal and fetal tissues, inclusion of a fetal compartment in the model adds little to the prediction of fetal methanol levels. The decision to use the non-pregnant model is supported by the data, and well justified in the narrative.
Sweeney	The extrapolation is adequate.

A4. EPA assumes limited methanol metabolism in the fetus because of limited alcohol dehydrogenase (ADH) activity in the human fetus, limited catalase and ADH activity in fetal rodents, and existing pharmacokinetic data that show nearly equal concentrations in maternal blood vs. the fetal compartment. Please comment on the validity of this assumption given the lack of data regarding potential alternate metabolic pathways in the fetus.

Reviewer	Comments
Burbacher	This seems like a reasonable assumption given the limited data available.
Byczkowski	Assumption of limited methanol metabolism in the fetus is consistent with the demonstrated low fetal metabolic activity for several other xenobiotics similar to methanol. As explained in the document P. B-3; L# 5 to 10: "...*The fact that measured fetal blood levels are virtually identical to maternal levels for methanol (and ethanol) tells us that the rate of metabolism in the fetus is not sufficient to significantly reduce the fetal concentration versus maternal*..." So, even though the contribution of potential alternate metabolic pathways in the fetus remains uncertain, this reviewer agrees with the simplified assumption, as explained in the document (P. B-2; L#19 to 21): "...*Because the maternal blood:fetal blood partition coefficients were near 1, there was no need to explicitly model fetal kinetics; they will be equivalent to maternal blood kinetics*..." Obviously, this assumption holds only if the parent compound (methanol itself) is indeed responsible for deleterious effects in the fetal/postnatal brain. As far as this reviewer is aware, to date, no convincing study explaining potential teratogenic mechanism of action of methanol was published.
Dorman	This assumption appears to be valid based on some of the existing methanol pharmacokinetic data (where maternal and fetal methanol concentrations are assessed). However, recent data published by Miller and Wells (2011) demonstrates that the embryotoxicity of methanol in cultured mouse fetuses is influenced by fetal catalase activity. In this study, methanol was more embryopathic in acatalasemic (aCat) mouse embryosthan their wildtype controls, with reduced anterior neuropore closure and head length only in catalase-deficient embryos. In concert with work published by Sweeting et al (2011) draw into question whether fetal methanol concentrations are a good predictor of teratogenic responses in different species – this data begs the question of which animal model(s) should be used in the methanol risk assessment where reproductive outcomes are of concern.
McMartin	The assumption of limited methanol metabolism in the fetus is probably justified based on the existing studies showing low levels of ADH and catalase in fetal tissues. However, these studies have technically measured these proteins using indirect measures such as immunoblotting showing protein amounts or activity measures with ethanol as the substrate. Ideally an activity measurement using methanol as the substrate would be needed to confirm the low activity of methanol metabolism in fetal tissues. Nevertheless, the assumption that the fetal tissues do not substantially impact the pharmacokinetics of methanol is likely a good assumption.
Roberts	The assumption of limited methanol metabolism in the fetus is valid from the perspective of PBPK modeling of methanol concentrations. There is sufficient information to show that ADH and catalase metabolism of methanol are relatively low in both the rodent and human fetus. Significant alternative pathways of metabolism of methanol in the fetus

	have not been identified. That is not to say that fetal metabolism is insignificant from the standpoint of methanol developmental toxicity, however. Pathways in the fetus that are quantitatively minor compared with maternal metabolism can nonetheless be very important in determining adverse effects. For example, studies by Wells and Miller (2011) suggest that fetal catalase activity is important in determining susceptibility to methanol developmental effects in rodents.
Salmon	It is obviously difficult to get objective measures of fetal metabolic capabilities, and although there are some data in this case there is inevitably a substantial measure of uncertainty involved. Some data described in the report (including the recent studies by Sweeting et al., 2010; 2011included in the addendum) suggest that although ADH1 levels in the fetus are typically low, this shortfall may be compensated by relatively higher catalase activity even in species where the latter route is not an important factor in the adult. However, the limited data available on methanol levels in the fetus relative to the maternal levels imply that fetal metabolic clearance is not sufficient relative to the rate of equilibration to substantially differentiate these two compartments, supporting the PBPK analysis described in the report.
Sweeney	The assumption appears valid.

A5. Please comment on the scientific justification of the extrapolation approach from rats to humans for in-utero and neonatal lactational and inhalation exposures.

Reviewer	Comments
Burbacher	As I mentioned above, the reviews of PBPK models for mouse, rat, and humans are comprehensive but it is not clear to this reviewer what the strengths and weaknesses are for each model and why the nonhuman primate model was not included in the final model development. Some clarification of the process for evaluating the usefulness of each model for the assessment and why the nonhuman primate model was not included would be helpful. There are also issues related to using the NEDO studies which included neonatal exposures that continue to be problematic, given the lack of data on lactational and early postnatal inhalation exposure to methanol.
Byczkowski	As explained in the section B.2.9 of the document (P. B-39, L#28 to P. B-41, L#2): "...Although the developmental endpoints of concern are effects which occur during in utero and (to a lesser extent) lactational exposure, it is not necessary for a MeOH PBPK model to specifically describe pregnancy (i.e., specify a fetal/gestational/conceptus compartment) and lactation in order for it to provide better cross-species extrapolation of risk than default methods..." Due to its physicochemical properties, it was logical to postulate that the concentration ratio of methanol in maternal vs fetal compartments should be consistently close to the unity, across different species. This allowed the modelers to simplify further the PBPK model and to perform interspecies extrapolations under assumption that the maternal compartment realistically represents the fetal one. Again, this assumption holds only if the parent compound (methanol itself) is indeed responsible for deleterious effect in the fetal/postnatal brain.

	The assumption of 1 to 1 ratio do not necessarily holds for lactational transfer of methanol, but it has been explained in the document (P.B-40, L# 18 to 36 and P. B-41, L# 1 to 2) that: "... *While lactational exposure is less direct than fetal exposure and blood or target-tissue levels in the breast-feeding infant or pup are likely to differ more from maternal levels, the health-effects data indicate that most of the effects of concern are due to fetal exposure, with only a small influence due to postbirth exposures*..." Whereas it would be prudent to add the lactatonal transfer of methanol in to PBPK model, but as discussed in the answer to A1 (above), for the purpose of risk assessment this may be not necessary.
Dorman	The use of the PBPK model is scientifically defensible and is a considerable strength of using a PBPK model. However, a concern remains as to whether the use of rodent data is appropriate since the metabolic activities and elimination kinetics of methanol in rodents are quite different from that seen in primates. As mentioned earlier, the model should be modified to include gestational and lactational components.
McMartin	Despite the various caveats noted above (A1-A4), the extrapolation conducted in this assessment from rats to humans is as good as can be done at present.
Roberts	As explained in Appendix C, experimental data indicate that the kinetics of inhaled methanol are similar in pregnant and non-pregnant mice, and that maternal blood and fetal methanol concentrations are approximately equal. This is assumed to apply to rats as well, and that maternal blood methanol concentrations are therefore appropriate indicators of fetal concentrations during gestation in this species as well. Because offspring of maternally exposed rats have subsequent exposure through lactation and direct inhalation, methanol concentrations in pups were likely greater than in the dams. Assuming that the same difference occurs in human mothers and offspring, this difference was deemed relatively inconsequential for the purposes of the analysis. The first assumption – similarity in maternal-fetal concentrations in both mice and rats – seems reasonable given similarities in the nature of distribution of methanol in the body across species. The second assumption requires a much greater leap of faith – that lactational and inhalation exposure postpartum in rats and humans are sufficiently similar that the same maternal/offspring methanol concentration ratios will be seen. The Review points out a study by Stern et al. (1996) indicating that methanol concentrations in pups exposed by inhalation were approximately 2.25 times higher than the dams, and states that a similar ratio probably occurred in the NEDO (1987) study used to generate the RfC, given similar designs of the studies. It is then assumed that the maternal/infant methanol concentration ratio in human infants would not be significantly greater than that observed in rats. This is purely speculation. Differences in exposure as well as methanol clearance between human infants and pups could lead to substantially different maternal/offspring methanol concentration ratios (higher or lower). This is a significant source of uncertainty in the extrapolation from rats to humans.
Salmon	Appropriate species-specific PBPK models have been developed, and validated against several independent data sets in each species. Although the human model is subject to some uncertainty at higher doses, it has been effectively validated for doses in the range of interest for development of the RfC and RfD. Given the data supporting relatively even distribution between mother and fetus, the extrapolation for exposures *in utero* appears to be well justified. The extrapolation for postnatal exposures, especially

	lactational, is subject to greater uncertainty, if only because of the diversity of feeding regimes and behaviors in human infants, but as noted in the report this exposure period is probably of less concern than the period *in utero* for sensitivity to the effects *in utero*deemed critical in determining the RfC and RfD. However, the model is justified for calculating exposures in infants with extensive lactational exposure, and these are certainly of concern in terms of health impacts later in the developmental process, such as possible neurobehavioral impacts.
Sweeney	Conceptually, the rat POD extrapolation to the human HED and HEC appear to be acceptable; the rat POD in and of itself, however, appears questionable. The human model parameterization also needs to be further evaluated (see A1).

(B) Inhalation Reference Concentration (RfC) for Methanol

B1. A chronic RfC for methanol has been derived from a perinatal inhalation study of the effects from exposing rat dams and pups to methanol during gestation and lactation (NEDO, 1987). Reference values from mouse (Rogers et al., 1993) and monkey (Burbacher et al., 1999; 2004) developmental studies, were also derived and discussed, but were not chosen for the RfC. Please comment on whether the selection of the principal study has been scientifically justified.

Reviewer	Comments
Burbacher	After reviewing the previous panel comments of the NEDO study, I do not believe that this is the most appropriate study for derivation of the RfC. The review indicated that there continue to be questions regarding the procedures used in the NEDO study (in utero and postnatal exposures, litter effects, etc) that make it difficult to evaluate the study for RfC derivation. The discussion on page 5-10 regarding the complications that arise from using the NEDO study where exposure was both gestational and postnatal postulates a number of assumptions that are supported by little or no data. Data on lactational transfer and early postnatal inhalation exposures are limited. A note: On page 5-15 the document indicates that the monkey VDR effects are an example of prenatal and continuing postnatal exposure effects. This is not accurate since the monkeys were only exposed prenatally. The study by Rogers et al., 1993 would seem to be the most appropriate choice at this time. The study is scientifically sound and robust. Exposures are limited to the prenatal period and the outcomes are clear. While I do not agree with several comments regarding the Burbacher et al monkey study, 1999, 2004 (see comments below), I do agree that the monkey study should not be used for derivation of the RfC due to the lack of a dose-response function for the major effects.
Byczkowski	The selection of NEDO (1987) as the study principal in developing RfC, seems to be reasonable, even though the selection has been justified on practical/technical grounds rather than scientific (c.f. P.5-5, L# 10 to 24): "*...Taking into account the limitations of the studies available for quantification purposes, decreased brain weight at 6 weeks in male Sprague-Dawley rats exposed throughout gestation and the postnatal period (NEDO, 1987) was chosen as the critical effect for the purposes of this dose-response*

	assessment as it can be reliably quantified and represents both a sensitive organ system and a key period of development..."
Dorman	Concerns remain concerning the use of the rat perinatal methanol study to derive the RfC. The endpoint of concern in this study was a decrease in neonatal brain weight (an effect also seen in a gestational only exposure albeit at a higher exposure dose). This response has not been replicated in other studies. Moreover, the analysis provided by the NEDO authors showed a gender difference (effects seen in males but not female rats). Moreover, the NEDO study relied on multiple t-tests as opposed to a more appropriate use of an ANOVA to evaluate gender and treatment responses. It is this reviewers understanding that this concern may not be relevant to EPA since EPA performed an independent BMD analysis of the data and demonstrated statistically significant trends. This should be more explicitly stated by EPA to alleviate concerns about the NEDO study. The EPA did provide alternative RfC calculations which were appropriate.
McMartin	Based on the analysis provided in the document and the results of the three studies themselves, the inhalation study of NEDO would appear to be the most justifiable of the three choices as the principal study. One advantage of the 1 generation or 2 generation NEDO rat studies over the others is the nearly continual exposure (20-22 h per day depending on the study) represents the types of exposures relative to the RfC/RfD (i.e. the daily exposure over the lifetime), whereas the mouse and monkey exposures were more like an occupational situation. A negative aspect of the NEDO study is that the critical effect (decreased brain weight) has not been reported in other studies, nor were there any corroborating clinical or pathological observations of depressed CNS activity noted in the rats in the NEDO study; in contrast, the critical effects in the Roger study (cervical ribs and CNS abnormalities) have been reported in other studies.
Roberts	Among the three developmental studies considered, only the Burbacher et al. studies in monkeys were discussed in terms of limitations. [Note: A number of the limitations in the Burbacher et al. studies are overstated, such as the inclusion of wild caught monkeys and the influence of C-sections on results.] The explanation for choosing the RfC from the NEDO study over the RfC from the Rogers et al. study is simply that the value is lower. Choosing the more lower, more conservative value is a policy choice rather than one based upon the scientific strengths of the two studies.
Salmon	The Agency reviewed a substantial number of studies, including those mentioned here, and in those considered adequate from a methodological standpoint dose responses were analyzed for a number of individual data sets. Selection of the principal study was explained and is in accordance with the usual guidelines recommending use of the study with the best data quality (including, in this case, availability of a validated PBPK model) and greatest sensitivity. The F_1 male Sprague-Dawley rats from the study by NEDO (1987) fit these criteria: other analyzable data sets were used as supporting studies. However, the Agency may need to reevaluate the study selection with regard to sensitivity if modification of the PBPK analysis for SD rats significantly alters the relative sensitivity (based on HECs) of the rat, mouse and monkey studies.

Sweeney	The selection of the principal study is contingent upon the details of the dose-response analysis and determination of the HEC/HED. Comparison of the SD rat PBPK model output for inhalation exposure with the experimental data suggests that the rat POD could be on the order of 6-fold too low (see A1). A revised rat POD would be similar to the mouse values. Regarding the monkey study, I do not believe it provides convincing evidence of an effect, given the inconsistencies in dose-response, multiple comparisons, and the potential for unreliable identification of "effects" in small studies. Admittedly, these areas are generally outside of my particular expertise.

B2. Reduction of brain weight at 6 weeks postnatally as reported in the NEDO (1987) developmental rat study was selected as the critical effect. Please comment on whether the rationale for the selection of this critical effect has been scientifically justified. Please identify and provide the rationale for any other endpoints (e.g., other reproductive and developmental effects reported in mouse and monkey studies) that should be considered in the selection of the critical effect.

Reviewer	Comments
Burbacher	As mentioned above, the Rogers at al. 1993 study would be a more appropriate study to use for deriving the RfC, using increased incidence of cervical ribs as the critical effect.
Byczkowski	For justification of the critical effect and principal study, see the answer to B1 (above). This reviewer is not aware of any more appropriate end point.
Dorman	There remains a general lack of transparency in the selection of this critical study and this single time point as the point of departure – this reviewer has the impression that the selection criteria used by EPA to determine the "best" study was the one that led to the lowest RfC. Although this is an appropriate approach (precautionary principle) the NEDO study remains problematic. Concerns about the statistical methods used by NEDO were raised by the external peer reviewers that reviewed this document for EPA; however, this concern does not appear to have been considered by the US EPA. I am also concerned that the EPA did not consider the full database from the NEDO study. Again, they arbitrarily considered only one time point (6 weeks) solely because it yielded the lowest value. This approach weakens the potential statistical power for a response that appears stable over a wide range of time points (3 to 8 weeks). In some ways, the NEDO study is the weakest of the three options. Although brain weight was evaluated there was a lack of histological or functional follow-up for this response. The Burbacher study uses the most appropriate species (monkey) and examined a wide range of reproductive and neurotoxicologcal endpoints and significant pharmacokinetic data. The Rodgers study has undergone independent peer review, documents responses reported by other laboratories, and has quite robust group sizes.
McMartin	Reduction of brain weight in the NEDO rat study seems to be the critical effect – although the choice of 6 week data for the analysis is not scientifically justifiable (just because it produces the lowest BMD may be "standard procedure", but is not scientific). The biggest problem with this choice as the critical effect is that the statistical analysis of the brain weights used by NEDO is wrong and the risk assessment has not acknowledged

	this error. According to the 1987 report on this study, NEDO utilized multiple t-tests to compare the results between the various groups. Considering that there were four treatment groups (control and 3 levels of methanol exposure), the analysis should have involved an analysis of variance (ANOVA) to determine if there were an overall effect on brain weight, followed by an appropriate post-hoc test to examine for differences among groups (in order to test if one of the levels of methanol was significantly different from control). This document shows the NEDO data in Table 4-10 with the footnote that the statistical differences were calculated by the authors. A re-analysis using ANOVA would seem to be appropriate to ensure that the brain weights are significantly reduced by a methanol treatment, thus allowing for the subsequent BMD analysis. The 10-15% decrease in brain weight in males is probably ”significant”, so this might be a moot point, but a re-analysis is necessary to determine which time frame and which methanol level are used for the BMD analysis. This issue is of greater importance because the NEDO study did not report (examine for?) any corroborating information including effects on the brain such as clinical signs/symptoms or CNS pathology, as noted above.
Roberts	A significant reduction in brain weight on an absolute basis was reported in the NEDO (1987) developmental rat study. No abnormal brain histopathology or functional deficits were noted. A recent review of this study conducted for EPA raised questions about the statistical analysis and whether or not the small brain weight changes in fact are significant and represent an adverse effect (External Letter Peer Review of Reports Documenting Methanol Studies in Monkeys, Rats, and Mice Performed by the New Energy Development Organization (NEDO), Peer Reviewer Comments, June 16, 2009). In contrast, the Rogers et al. study is considered rigorous and well described (see, for example, NTP-CERHR, 2003), and the increases in cervical ribs and supernumerary ribs observed in this study could be considered a more scientifically justified critical effect.
Salmon	The report provides a thorough presentation of the various endpoints available for consideration from the NEDO (1987) studies and other reports. The various endpoints are reviewed for sensitivity, statistical significance and general data quality. These criteria and the standard risk assessment procedures, including selection of the most sensitive endpoint (subject to data quality considerations) point to the NEDO (1987) brain weight reduction data as the best choice for the critical effect. The other endpoints provide important supporting information, but are generally either less sensitive or less reliable (especially the primate studies which necessarily involve fewer maternal subjects per group and smaller litter sizes, as well as showing various deficiencies in the experiments as reported.). However, the Agency may need to reevaluate the endpoint selection with regard to sensitivity if modification of the PBPK analysis for SD rats significantly alters the relative sensitivity (based on HECs) of the rat, mouse and monkey studies. Standard guidelines indicate the choice of the most sensitive strain and endpoint available, subject to limitations of study design and quality.
Sweeney	As noted for B1, the discrepancies between the SD rat inhalation PBPK model and data lead me to question the internal dosimetry estimates that led to this study being identified as “critical”.

B3. Benchmark dose modeling of decreased pup brain weight relative to maternal internal methanol doses predicted by the PBPK model was used to derive the point of departure (POD) for the RfC. Has the BMD/PBPK approach been appropriately conducted? Has adequate justification been provided for the selected internal dose metric, i.e., area under the curve (AUC) for methanol, in the blood of dams? Please identify and provide the rationale for any alternative approaches for the determination of the POD, including choice of another dose metric (e.g., methanol metabolized), and discuss whether such approaches are preferred to EPA's approach.

Reviewer	Comments
Burbacher	This is not my area of expertise so I will not comment.
Byczkowski	The BMD/PBPK approach has been appropriately applied in the derivation of RfC, using health protective default (change of one S.D. from control mean) to determine BMR, according to the U.S. EPA guidance. The selection of AUC of methanol in maternal blood as a surrogate dose metric for dose-effect modeling of postnatal changes in brain also seems to be technically correct. Again, as discussed in the answer to A1 (above), the selection has been justified on practical/technical grounds. Without understanding of the exact mechanism of action of the chemical, selection of any surrogate dose metric is somehow speculative.
Dorman	BMD modeling: "No response as this is outside my area of expertise." The EPA could have considered either AUC or Cmax as the internal dosimetric of interest. The Agency has not adequately explained its rationale for the use of AUC rather than Cmax (e.g., see literature related to methanol and 2-methoxyethanol). I endorse the use of blood methanol as the dosimetric of interest. I am confused by the rationale used by EPA to calculate the AUC. Table 5-2 indicated that the AUC was calculated with a 5 day 22 hr/day simulation. Why was a 5 day exposure duration was used?
McMartin	The BMD/PBPK approach has been conducted appropriately for the most part (with the exception noted in B2). The selected internal dose metric (methanol AUC) is preferable over the alternative of methanol metabolized. Although there are studies showing that formaldehyde (HCHO) is the most "toxic" of the three compounds (methanol, formate and HCHO) in whole embryo cultures, this is a common situation with HCHO since it is a very interactive aldehyde. The key question is whether any of the metabolites are transported into the fetus or whether they are formed in the fetus (and if so, reach the fetal brain). Most of the existing studies have ruled out a role for formate in these developmental effects, so the only question is the potential role of formaldehyde. Certainly based on all existing studies, HCHO is not formed in the dam/mother and transported into the fetus. Also, fetal metabolism of methanol to HCHO appears to be minimal if at all. Even if a small amount of HCHO were generated in the fetal liver, it would not be transported to the brain (because it would be rapidly metabolized to formate within the fetal liver cell or would rapidly bind to components in the cell – either way it is not likely to even leave the liver cell).

	The increased effect seen when methanol is administered to glutathione (GSH)-depleted animals does not necessarily imply that the MOA for methanol involves metabolism to HCHO (which has been suggested because GSH depletion should decrease HCHO elimination allowing for higher HCHO levels). Depletion of GSH, as the major cellular antioxidant, will also increase the accumulation of reactive oxygen species (ROS) – since generation of ROS is a viable mechanism by which methanol induces its effects, the increase in ROS may explain the increased effects of methanol in GSH-depleted animals. The existing studies do indicate that methanol itself is the responsible agent (although possibly through generation of ROS by unknown mechanisms). Hence, the dose metric is either methanol AUC or Cmax. The justification for using AUC is tenuous – there are some studies suggesting that duration of exposure is important, but the fact that the effect on brain weight does not differ between the 3, 6 and 8 week periods does conflict with this assumption. Use of Cmax would suggest that there is a threshold response, i.e. a certain blood level must be reached. This alternative is attractive (formate Cmax is the key to ocular toxicity for example).
Roberts	An explanation is provided in Appendix C why maternal AUC is considered to be an appropriate dose metric for offspring. There are two issues: 1) use of a maternal dose metric to represent offspring; and 2) choice of AUC versus another metric for internal dose. The first issue is discussed in the response to Question A5. With respect to the second issue, internal dose metrics can be selected based upon mechanistic or empirical considerations. Given that the mode of action of methanol developmental effects is unknown, empirical evaluation is left. Unfortunately, there is little in the way of data for empirical evaluation. Because there is evidence that effects on brain weight undergo some recovery when exposure is terminated, this was viewed as indicating that both the level and duration of effect are important (see pg C-2). Because AUC incorporates a time component, and because it is commonly used as an internal dose metric, it was selected for this endpoint. AUC is a reasonable choice, providing a measure of the average concentration over time. Other metrics could be considered, but as noted above, there are no data with which to argue that any would be a better choice.
Salmon	Several possible choices for the dose metric (AUC, peak concentrations, amount of methanol metabolized) are discussed both in the context of their availability from the PBPK model and their possible relevance in the light of studies of the mechanism of action and time-course of exposure during sensitive periods for teratogenesis. Studies of possible modes of action for methanol toxicity have identified the importance of formate in acute toxicity, but imply that this is not a major element of the mechanism for the developmental effects. One possible mechanism for the impacts on the fetus is a direct effect of methanol itself, or of its proximal metabolite formaldehyde. Another possibility is the generation of reactive oxygen species (ROS) as a side effect oxidation of methanol. This hypothesis appears to be popular with some investigators, but actual evidence of its importance seems to be limited. In view of these mechanistic findings the choice of a dose metric based on methanol rather than downstream metabolites appears sensible. The report provides adequate justification for selection of the AUC for methanol as the appropriate internal dose metric in analyzing the NEDO (1987) data, which involve an ongoing exposure time element. (For the single-day exposure experiments by Rogers et al. [1995] the non-time-dependent dose metric C_{max} for methanol was shown to be more suitable, as described in Appendix D.). Even if the metabolism-related formation of ROS or formaldehyde are important contributors to the observed toxic effects, a methanol-based dose metric is applicable when the downstream metabolic processes

	such as removal of ROS or formaldehyde are much faster than the rate-limiting oxidation of methanol. Conduct of the BMD analysis was correct and in accordance with the usual guidance for this approach, and was thoroughly reported. The report is correct in noting (on page 5-15) that a 5% BMR is appropriate for analysis of quantal data from the developmental studies considered here: however, it is incorrect to comment as was done here that "a 10% BMR is adequate for moist traditional bioassays". Although this proposal was included in the original guidance for the BMD methodology, subsequent experience by various risk assessors (including Agency staff in a number of recent assessments) have concluded that the 5% BMR is more appropriate for identifying a POD to which the standard UFs are applied (i.e. no UF_L) in a standard animal study with quantal data. Neither the 5% nor the 10% BMR have any particular *a priori* justification for continuous data: the default assumption in this case is the BMR of 1 standard deviation of the control dataset (as preferred here). In any case the data need to be examined to determine an appropriate BMR representing a minimal detection level or threshold of biologically significant response: this especially applies for continuous data.
Sweeney	The dose metric (blood methanol) is appropriate, but does not appear to have been reliably computed for rats. The BMD approach appears to have otherwise been appropriately implemented.

B4. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the POD for the derivation of the RfC. It is assumed that these UFs account for variability in methanol dosimetry among human newborns following gestational and lactational exposure, and for uncertainty regarding the ratio of newborn-dose to maternal-dose in humans. Please comment on these assumptions and on the scientific justification for the selected UFs.

Reviewer	Comments
Burbacher	The rationale for selecting the uncertainty factors is consistent with EPA policy.
Byczkowski	The following uncertainty factors, totaling UF of 100-fold, were applied in the development of RfC: UF_H = 10; UF_A = 3; and UF_D = 3. Even though, selection of uncertainty factors is usually arbitrary to some extent, the justification provided in the document seems to be convincing and in compliance with the health-protective U.S. EPA guideline.
Dorman	The UFs are poorly justified by the EPA. This is magnified when one reviews the alternative RfCs derived by EPA (Table 5-4)/ For example, it is surprising that the EPA used the same interspecies UF for rodent and nonhuman primate studies – given the fact that significant species difference exist between rodents and humans and less so between monkeys and people (use UF = 1). The database UF is poorly justified – the EPA continues to use this UF even in the face of a very rich toxicology database. When, if ever, will the database be adequate? The EPA is also basing the POD on the most sensitive studies that were conducted in neonatal animals from multiple species. If the critical study used is in neonates then why is an additional UF of 3 needed to account for

	children as a susceptible subpopulation. Again these UFs might be justified but they are poorly justified. Likewise, the discussion of possible gender effects (page 5-19) doesn't clearly articulate whether the experimental designs AND statistical analyses were adequate to determine whether or not a gender difference existed. Just as importantly, the EPA systematically chose the most conservative approaches in developing the RfC. For example, they used a single SD for BMDL rather than a 4 or 10% changes as commonly used in some noncancer risk assessments (e.g., see page 5-23). These decisions were driven by the goal of obtaining a lower RfC value. There is a lack of transparency in this approach. In aggregate this yields an RfC value that
McMartin	The 10-fold uncertainty factor for inter-human variability is generally justified by likely genetic variations in methanol metabolizing enzymes (thus, producing differences in methanol elimination) and by highly likely differences in human nutrition (folate deficiency is known to exacerbate the developmental effects). The 3-fold uncertainty factor for the pharmacodynamic human-to-animal extrapolation is also likely justified based on the lack of existing knowledge regarding the mechanism of the developmental effects – hence not enough information to assess pharmacodynamic differences in animals vs. humans. The database uncertainty factor of 3 is not at all justified. Based on standard procedures, there is never enough data to be certain regarding a risk "assessment" (that is why it is called risk assessment not a risk determination). More importantly, conservative assumptions are always used on all these procedures (PBPK assumes the most conservative scenarios, BMD analysis itself favors the conservative numbers and lastly when given the choice of alternative BMD numbers such as those obtained from the 3 vs. 6 vs. 8 week data, the lowest, i.e. most conservative, number is chosen. Thus, because of the conservative approach in this risk assessment and because of the fact that methanol has an endogenous profile, there is no reason for the additional UFd of 3.
Roberts	A composite UF of 100 was selected based upon a UF_H of 10, a UF_A of 3, and a Database UF of 3. A UF_A of 3 was selected because PBPK modeled was thought to address the pharmacokinetic component of this UF. A Database UF of 3 was chosen based upon perceived deficiencies in the toxicity database, mentioning uncertainty regarding the most appropriate test species and limitations in developmental toxicity studies currently available. Selection of individual UFs appears to be consistent with contemporary EPA guidance and practice, although a strong argument could be made for eliminating the database uncertainty factor. As noted in the question, it is implied that the UFs cover specific uncertainties in derivation of the RfC, including variability in methanol dosimetry among human newborns following gestational and lactational exposure and the ratio of newborn dose to maternal dose in humans. Because of limited information available, it is difficult to judge the potential error associated with assumptions regarding human newborn exposure and resulting methanol concentrations. However, it seems likely that the error is not so great that a composite UF of 100 would be inadequate.
Salmon	The report identifies the potential for significant inter-individual variability in the effects on humans to methanol, including sensitive sub-populations with various enzyme polymorphisms, folate deficiency and enhanced exposure and sensitivity of infants and children. However there does not appear to have been much effort to actually quantify this variability (perhaps due to insufficient data), and the standard default value of UFH

	= 10 is used. It is a matter of some debate whether this factor is in fact sufficient in the general case, although the ranges of variation in enzyme levels suggested do not appear to be as large as for some other toxicants. Some further examination and discussion of this issue would be helpful in establishing the limits of the data available to inform the decision on the value for UFH. The use of a UF of 3 to account for the uncertainty in the toxicodynamic component of the extrapolation from rodents to humans is an appropriate application of the default, since no independent data are available to further quantify this extrapolation. Use of a value of 1 for the toxicokinetic component of this uncertainty is also appropriate in view of the use of a comprehensive and validated PBPK model for both humans and the rodent test species. The specific issues of toxicokinetic uncertainty for gestational exposure, and inhalation or lactational exposure of newborns, are discussed in the document. The question of fetal vs. maternal exposure appears to be relatively well addressed in animal models, and there is no reason to expect major differences for humans since the explanation rests largely on chemistry of the toxicant, rather than the physiology or anatomy of the subject. It therefore appears that this question is not a major source of uncertainty. The exposure and toxicokinetics of the newborn does appear to involve somewhat greater uncertainty, although some of this is associated with differences in diet and behavior and is probably accommodated by the allowance of a UF of 10 for human inter-individual variability. There is no particular reason to expect greater uncertainty in this respect with regard to the interspecies extrapolation. Use of a factor of 3 for database uncertainty is justified by the absence of rigorous developmental neurotoxicity tests in rodents. Such effects could reasonably be anticipated given the suggestive (but quantitatively inconclusive) results in monkeys, and the clear observation of anatomical impacts on the central nervous system of rodents. This is a source of concern since changes in brain weight imply a relatively substantial change in CNS development which is quite likely to have functional impacts, and in many cases primates have proved to be more susceptible to these effects than rodents. It appears that the results of the developmental study in primates (Burbacher et al, 1999) support this concern. It seems unlikely that this latter study could be used as the critical study for determination of an RfC due to its design limitations, but if it were these limitations would in any case need to be represented by an appropriate UF. The Agency applied the UFs to the POD expressed as a HEC, which is the standard procedure and is preferred to alternative suggestions that the UFs be applied to intermediate measures such as blood concentrations or AUCs.
Sweeney	The uncertainty factors UFH and UFD are inadequately justified. At the level of the proposed RfC and RfD, intraspecies differences in disposition of exogenous methanol in humans will likely have no meaningful impact on the body burden of "total" methanol. Thus a full UFH of 10 is not warranted. Sensitivity analyses of the human PBPK model for methanol could identify parameters which have an impact on total or "additional" methanol, but were not conducted. Identification of the potential impact of variability and/or uncertainty of the human model parameters on predicted body burden of methanol could further inform the selection of UFH values. The discussion of UFH inappropriately includes the special sensitivity of children. Since the database includes two-generation studies (in fact, the current key study is a two-generation study), there is no reason discuss children's potential susceptibility; no

	particular developmental susceptibility of humans vs. test species is expected. The authors appear to be attempting to double dip on an uncertainty factor that is not needed. It is hard to imagine how a UFD of 3 can be justified. As the authors note, the key endpoint is developmental toxicity, which has been evaluated in multiple species, including primate, and special endpoints such as neurotoxicity and immunotoxicity have been evaluated. There is no need to have a UF because "there is uncertainty regarding which test species is most relevant to humans"—the lowest, high-quality point of departure was used. There is also no need to have a UFD for "dose spacing" because the BMD analysis counters this potential design deficiency. Also, I personally believe it is most appropriate to apply the uncertainty factors to the internal dose point of departure, prior to interspecies extrapolation with the pharmacokinetic model to account for non-linearities in external vs. internal dose relationships. EPA should discuss their choice of applying UFs to the HED rather than the BMDL. While this point is moot when the kinetics are linear with respect to exposure concentration, this is not the case with methanol. In the case of EPA's March 2011 analysis for the rat, the difference appears to be small (RfC of 1.8 mg/m^3 using HEC/UF, 2.0 mg/m^3 using BMDL/UF, based on my calculations, ~10%), but the difference becomeslarger starting from a higherpoint of departure. For example, if UFs are applied first to the mouse cervical rib BMDL05, then converted to the candidate RfC using the PBPK model, the candidate RfC increases from 10.4 mg/m^3 to 23.6 mg/m^3 (>2-fold).

(C) Oral Reference Dose (RfD) for Methanol

C1. EPA concluded that the oral RfD should be derived using a route-to-route extrapolation from the more extensive inhalation database given the paucity of oral toxicity data. Please comment on whether the rationale for this approach has been scientifically justified and clearly explained. Please identify and provide the rationale for any alternative approaches for the determining the RfD and discuss whether such approaches are preferred to EPA's approach.

Reviewer	Comments
Burbacher	This is not my area of expertise so I will not comment.
Byczkowski	The selection of route-to-route extrapolation from inhalation study to develop oral reference value is clearly explained in the document (e.g. on P.5-23 L# 11 to 15): "...*The limited data for oral administration indicate similar effects as reported via inhalation exposure (e.g., the brain and fetal skeletal system are targets of toxicity). Methanol has been shown to be rapidly and well-absorbed by both the oral and inhalation routes of exposure (CERHR, 2004; Kavet & Nauss, 1990). Once absorbed, methanol distributes rapidly to all organs and tissues according to water content, regardless of route of exposure*..." This reviewer is not aware of any more appropriate approach and/or study that would be relevant to human chronic oral exposure to methanol. However, this is remarkable, that the RfD derived in this document is numerically almost the same as the previous one, derived from rat oral subchronic study.

Dorman	Theoretically this approach appears appropriate since there is limited data to consider that the route of exposure influences methanol disposition once this alcohol is absorbed. The US EPA should provide alternative RfC estimates that would be derived using traditional approaches.
McMartin	The existing oral studies do not appear to be adequate for determination of an RfD. Either no dose-response effects were reported in the studies (Soffritti) or the data were not suitable for BMD determination (EPA 86). Because of the thoroughness of the PBPK model as it is currently developed, the route-to-route extrapolation from inhalation data to estimate the oral RfD is a logical approach. The explanation and justification is adequate.
Roberts	Limitations in existing oral studies are adequately described in Section 5.2.1, supporting a decision not to use oral study data to develop an oral RfD. Observations of effects in subchronic and chronic oral toxicity studies by U.S. EPA (1986) and Soffritti et al. (2002), respectively, were insufficient to support quantitative analysis to establish a NOAEL or BMD. Also, I agree with the decision not to use any of the oral methanol reproductive and developmental toxicity studies that used very high methanol doses over selected gestational periods. Parenthetically, I am surprised that better oral chronic toxicity data are not available for a chemical of the importance of methanol. Under the circumstances, the decision to use route-to-route extrapolation and inhalation toxicity data is justified, and the rationale is clearly explained in the Toxicological Review (pg 5-23). I have no suggestions for an alternative approach to developing an oral RfD that would be preferable.
Salmon	This choice was clearly explained and justified in the report. Given the availability of a detailed PBPK model which has been validated for several species and routes of exposure, it is entirely reasonable to use this model for route-to-route extrapolation from the more extensive and reliable inhalation data to the oral route. Although oral developmental studies are available which suggest similar developmental impacts to those seen by inhalation, these studies used relatively high doses and were therefore not suitable for determining a minimal effect POD on which the RfD could be based. Route-to-route extrapolation is a well established procedure in risk assessment which has been used on a number of previous occasions when appropriate data are available to support it. This is clearly preferable to the alternatives of either developing an RfD based on route-specific but inadequate data, or not developing an oral standard at all.
Sweeney	Route-to-route extrapolation appears to be justified. Human model validation using the oral data of Schmutte et al. (1988) (see D.2) would further strengthen confidence in the route-to-route extrapolation.

C2. A PBPK model was used to derive the RfD via a route-to-route extrapolation, in which the internal-dose POD used for the derivation of the RfC based on data from the NEDO (1987) study was extrapolated to human oral exposure levels using the human PBPK model. Please comment on whether the rationale for this approach has been scientifically justified. Has adequate justification been provided for the selected internal dose metric, i.e., AUC for methanol, in the blood of dams? Is the PBPK model suitable for extrapolation of fetal and neonatal endpoints to human oral exposures? Please provide a detailed explanation.

Reviewer	Comments
Burbacher	This is not my area of expertise so I will not comment.
Byczkowski	The reviewer's comments in answers to B1 through B3 (above) regarding RfC, are relevant also to this question about RfD. The rationale provided in the document for use of route-to-route extrapolation has been quoted in the answer to C1 (above). Regarding the AUC for methanol in blood of dams, selected as an internal dose metric, again - "*without understanding of the exact mechanism of action of the chemical, selection of any surrogate dose metric is somehow speculative*" (quoted from answer to B3, above).
Dorman	As mentioned earlier, the model should include gestational and lactational components. This remains a weakness of the EPA approach.
McMartin	Despite the comments on the PBPK model in Section A above, the developed model is useful for the extrapolation from the NEDO inhalation study to produce an oral POD. As in C1, the rationale and justification is adequate. Despite the various caveats noted in Section B, methanol AUC is suitable.
Roberts	The route-to-route extrapolation is based on assumptions that the critical effect is related to methanol concentrations in blood and is independent of the route of exposure by which methanol reaches the blood. Both assumptions are reasonable. It follows that the POD and the dose metric should be the same whether from oral or inhalation exposure. The PBPK model is suitable to extrapolate an external human oral dose that corresponds to the internal dose POD. The appropriateness of the POD with respect to fetal and neonatal endpoints from oral exposure is the same as with inhalation exposure, i.e., has the same limitations (see answers to previous questions).
Salmon	The development of the PBPK model and identification of a POD for the internal dose metric is essentially the same for both inhalation and oral exposures, and has been well described and justified in the report. Similarly, the considerations for selection of the internal dose metric (AUC for methanol) and questions about whether the maternal AUC is reflective of the critical concentrations in the fetus are largely independent of the route. It appears since methanol is readily distributed throughout the aqueous compartments of the body, and that rates of metabolism in the fetus are not so great as to significantly perturb this close-to-equilibrium distribution. Maternal blood concentrations of methanol are therefore expected to be adequate predictors of the concentrations in critical tissues of the fetus for either inhalation or oral exposures. The main difference for the oral route is the discontinuous exposure pattern in rodents (and also humans, although this is harder to define for the general case), and the finite holding time in the stomach

	prior to absorption into the systemic circulation. The description of the model includes features to address these points, including efforts to validate the model against independent oral data. Earlier comments (Section A1) as to the possible re-interpretation of model parameters for the SD rat apply for the oral model as well as the inhalation model.
Sweeney	With the exception of numerical value of the point of departure and the order in which UFs are applied/extrapolations made, the approach appears to generally be valid. Maternal blood is an adequate surrogate for fetal dosimetry, based on the available data.

C3. EPA applied the same UFs to the POD for the derivation of the RfD as for the RfC.Please comment on the rationale for the selection of the UFs.

Reviewer	Comments
Burbacher	Again, the rationale for selecting the uncertainty factors is consistent with EPA policy. Although I would assume there would be more uncertainty due to the lack of oral exposure data.
Byczkowski	The health-protective UF totaling 100-fold, the same as in developing RfC, has been applied to RfD. As I stated in the answer to B4 (above): "*Even though, selection of uncertainty factors is usually arbitrary to some extent, the justification provided in the document seems to be convincing and in compliance with the health-protective U.S. EPA guideline*".
Dorman	The database for oral and inhalation are very different. Using an identical database UF is unexpected.
McMartin	The same comments regarding UFs in section B4 can be applied here. Just to reiterate, the UFd of 3 is not justifiable and should be eliminated.
Roberts	The Review states that because the same dataset, endpoint, and PBPK model were used to derive the RfC and RfD, the same UFs were applied. I agree with this rationale. As noted in the response to B4, it is questionable whether a database uncertainty factor of 3 is needed given the data available on sensitive endpoints in multiple species.
Salmon	Given the route-specific calculation of oral HED using the PBPK model, and the use of the same underlying toxicity database and critical study, the justification for the UFs is the same as that noted in the discussion of the RfC derivation. There may be some minor differences in the extent of interindividual variation in exposure or susceptibility between routes, but these are probably not sufficient, or sufficiently well characterized, to justify any change in the chosen value of UF_H. It may be argued that the reliance on route-to-route extrapolation involves some additional uncertainty, but given the use of a validated model, and the observation of similar developmental effects in oral studies judged adequate for hazard identification but not for dose-response assessment, there does not

	seem to be a strong case for changing either UF_A or UF_H on this account. The arguments in favor of a value of 3 for UF_D are exactly the same regardless of route.
Sweeney	My comments above (question B4) apply equally to the UFs for the RfD.

(D) General Charge Questions

D1. Is the Toxicological Review logical, clear and concise?Has EPA clearly synthesized the scientific evidence for non-cancer hazards?

Reviewer	Comments
Burbacher	There is probably not much that can be done about it since I assume there is a format for these reports, but I found the assessment to be very repetitive. I would like to commend the authors of the assessment for a very comprehensive review. The only comment I have aside from the repetitive nature of the review, is that it may be good to provide some discussion and/or figure related to "decision trees" for major decisions that were made in terms of choices of approaches, studies, outcomes, etc. to make these choices more transparent than they currently are. The comments regarding the Burbacher et al study seemed inconsistent and at time inaccurate (see below for specific comments regarding the Burbacher et al study).
Byczkowski	The reviewed Toxicological Review document is logical and mostly clear. However definitely it is not concise. It provides not only the synthesis of evidence for noncancer health hazards, but it also attempts to describe most of the extensive research related to methanol toxicity, sometimes in a redundant manner. The length and the redundancy of the document are, perhaps, resulting from the formal structure required for IRIS Toxicological Reviews by U.S. EPA. It seems that in the revision of the current document, the U.S. EPA could use the recent NAS recommendation that "*EPA should edit documents to reduce the text volume and address redundancies and inconsistencies*".
Dorman	The document has a number of significant weaknesses. Many of these were raised with the draft IRIS assessment for formaldehyde (NAS 2011). These include: • Extensive reliance on a narrative approach with significant repetition of information throughout the document. • There is a lack of inclusion and exclusion criteria. For example, several studies that could be supportive are lacking from the document – e.g., Bolon et al (1993, 1994) There are also other studies, including work in monkeys, with aspartame that may be supportive (e.g., Reynolds). Since Table 3-2 includes results from aspartame exposure this does not seem to be a clear exclusion criterion. Likewise, search terms and databases examined have been poorly defined. • The document becomes extremely speculative. One of the more problematic section(s) related to the possibility that formaldehyde is the teratogen involved.

	Granted formaldehyde is more embryotoxic than either methanol or formate (on an equi-molar basis) this finding may or may not be related to the spectrum of teratogenic responses seen in animals (e.g., cervical rib anomalies, exencephaly, etc). This is especially important since there is a paucity of in vivo data suggesting that formaldehyde is teratogenic in animals (or people) (NAS, 2011). Another example relates to the discussion of parkinsonian signs in methanol-exposed people. These effects are commonly seen in hypoxic brain conditions so the question remains whether or not these represent a primary response or a secondary effect. • The US EPA should rely more heavily on tables and not replicate repeatedly the same description of individual research studies. • The Appendices are extremely difficult to read – there is an enormous amount of extraneous information provided. • Table 3-3 should include the Dorman cynomolgus monkey study with a clear indication that it involved lung only exposure of anesthetized monkeys. • In multiple locations the EPA mentions the results of the Fagan test performed by Burbacher and coworkers in monkeys exposed perinatally to inhaled methanol. At times, these results are used to support the selection of the NEDO rat study as the critical study. However, this effect lacked statistical significance in these animals. Is the EPA concluding that the effects seen were biologically significant despite the lack of a statistically significant response? • I found section 3.4.2.4 confusing. There are other models that have been developed (Yoon et al., 2010, 2011) with inhaled manganese that could form the basis for a gestaional and lactational model. • The discussion of a two compartment stomach (page 3-28 and elsewhere) for rodents need additional justification (squamous and epithelial portions?). Is this structure appropriate for people (as indicated on page 3-51). • The EPA uses terms that describe model fits as “quite poor” (e/g/. see page 3-40 and elsewhere). This is at best a qualitative term that needs to be better clarified (visual inspection, goodness of fit, other?). • I also found the use of different units of measure (e.g., mg/dL, mM, or ppm versus mg/m^3) frustrating. It would be ideal to pick one set of units (ppm would be preferred until calculation of the actual RfC value). • Page 4-7 (and possibly earlier) the use of alcohol dehydrogenase inhibitors as a clinical ‘antidote’ should be discussed here. Many readers may not be familiar with this treatment approach. • Page 4-18 – change to uriniferous tubule. • Page 4-40 – does folate deficiency affect methanol concentrations significantly? Which data support this conclusion?

	• Page 4-48 – change combustion product to metabolite. • Table Legends should include exposure duration as well as exposure concentrations.
McMartin	In general the Toxicological Review is thorough in its description of the numerous studies as well as of the BMD and PBPK modeling. Two general comments – 1) it is somewhat repetitious, which makes it difficult to appreciate the key points (having to wade through the extraneous material also); 2) the inclusion of the discussion of the CNS effects produced by acute methanol overdosing is not appropriate – it seems in a biased way to validate the subsequent choice of the NEDO study (decrease in brain weights suggesting a methanol-induced CNS effect). In reality, the two effects are not related (the acute methanol CNS pathology occurs in the exposed subject per se, due most likely to formate accumulation, while the reduced brain weight is a developmental effect unrelated to formate). The following minor comments on the document are noted. 1. Table 3-5 – why are the ranges of blood methanol for the Stanton study expressed downwards (from high to low), while the other ranges go from low to high? IS it true that the blood formate values reported by Horton are identical for the F-344 rats as for the monkeys (Table 3-3)? The last entry in Table 3-5 for Lee et al is not clear – the numbers for methanol and formate do not align with the various dose groups so it is hard to tell which is which. 2. Section 3.4.2.1 is identical to section 3.4.2.3 – one can be deleted. 3. Table 3-11 and p. 3-40. The data are cited as mean $\pm$ SE in text but as SD in the table. 4. p. 4-2. The reports by Bennett et al. and by Benton/Calhoun are really from the same study of the same epidemic. This should be clarified here (or one reference eliminated). 5. p. 4-6. As written, it sounds like there have only been two cases of inhalation/skin exposure. Not true, there have been many published reports, although mostly by Woods and coworkers in the early 1900s. 6. p. 4-14. I realize the data quoted come from an encyclopedic listing of LD50s, but as listed, the data imply that the methanol is not very lethal to the monkey. IN fact, monkeys are quite susceptible to the acute oral toxicity of methanol, with lethality occurring in the 3-4 g/kg range which is much less than in the rodent. A caveat should be added to indicate the unique sensitivity of the primate to methanol. 7. p. 4-70 – it should be Table 9 not 10 in the first line. Also in line 19 the referencing is misplaced. Johlin et al studied hepatic folate levels between species and not anything to do with catalase levels (so reference should be in the previous sentence) 8. p. 83 line 19, Latter not former.

Roberts	The Toxicological Review is generally logical and clear, but not concise. The format of the document contributes to redundancies, and presentation of some topics is fragmented, forcing the reader to synthesize information presented in more than one section of the main document and appendices. This is problem inherent in the current format for IRIS toxicological reviews. Sometimes subtle points are lost in the repetition. A great deal of information in order for the analysis to be transparent, but this shouldn't get in the way of clearly highlighted key points and decisions. A different format could be much more effective in conveying critical information, interpretations, and decisions regarding available, relevant toxicological literature.
Salmon	The review is thorough and well written, and takes care to provide descriptions of the available evidence in a clear, complete and unbiased form. The report presents a careful and well justified synthesis of these data. The decision to review the toxicokinetics in the first section of the report is a departure from the usual format, but in view of the importance of this topic for methanol and the extent of this section this is a defensible choice.
Sweeney	<u>Major comments</u> EPA has inadequately synthesized the SD rat toxicokinetic data in their PBPK model (see A.1). While chapter 3 (Toxicokinetics) and Appendix B (PBPK model) cover many of the same data sets, it is clear that they were written by different people and that no one took the time to ensure that the reader can seamlessly move between these sections, as one would expect in a more cohesive document. For example, key studies from the PBPK modeling are omitted from Chapter 3 summary tables 3-2, 3-4, and 3-5 (Ernstgard et al., 2005; Rogers et al., 1997, Ward et al., 1997). While the same data may have been presented in the studies that are listed, it is much harder to follow as is. Also, relevant studies noted in chapter 3 were not used in model development/validation (see A1). The clarity of the document is hampered by the lack of a clear synthesis of evidence regarding plausible modes of action for developmental toxicity. <u>Minor comments</u> Page 3-3. This table (from CERHR 2004) is not up to date, omitting the data of Ernstgard et al., 2005. Page 3-5. The reported blood methanol levels for the NEDO studies are unclear. Are the "0" ppm data the background methanol levels, and the 10, 100, and 1000 ppm entries measured values minus background? Page 3-13. Methanol blood data for monkeys are discussed, but not shown, while formate and folate data are shown. Considering EPA proposes that the MOA and key dose metric are related to methanol, not the metabolites, it seems that the methanol data should be shown, and the formate and folate data are less important. Page 3-25, Table 3-9. This table is not as informative or useful as it could be. Key studies are omitted (e.g., Ernstgard et al., 2005). Organizing the data by alphabetizing the first authors is less helpful than grouping by species and strain. Some of these data

	were not used for the EPA-developed PBPK model (e.g., Gonzalez-Quevedo et al., 2002). In some places, animals are described only as pregnant, while in others, Gestation Days are clearly specified. It is not clear where duplicate reporting of same data has occurred (e.g., Perkins et al., 1995a, 1995b, 1996 and Pollack and Brouwer, 1996, Pollack et al. 1993). Page 5-11, Table 5-2. Clarity would be improved by converting steady state daily AUC to time-weighted average concentration. These concentrations could more readily be compared to the experimental data and are more readily understood by non-experts. Page 5-14, Figure 5-1. The figure caption should provide the units for the response and dose; better yet, this information should be on the figure's axes. This comment applies equally to all of Appendix C. Page B-8. The metabolic parameters for the mouse do not correspond to "one high affinity/low capacity and one low-affinity, high-capacity enzyme". The lower affinity enzyme, as described by Vmaxc2 and KM2, is also much lower capacity (3.2 vs. 19). Page B-38. Daily AUC can and should be converted to TWA concentration as noted above (comment on page 5-11). Page B-40. At this point, the authors introduced the reference Pollack and Brouwer, 1996, which appears to cover data published elsewhere (e.g., Perkins et al.). This was initially confusing, because I thought new data were being introduced. (I read Appendix B prior to reading Chapter 3.) Page C-42, line 11. It is incorrectly stated that the monkey BMD modeling was done on the basis of external concentration (ppm). Page C-44, Table C-10, footnote a. It is incorrectly stated that the AUC for the monkey study was estimated using a rat PBPK model. <u>Nit-picky comments (e.g., typographical errors, unclear referencing.)</u> Pages 3-20, 3-21. The text of section 3.4.2.1 is identical to the text of section 3.4.2.3. Page 4-6, line 3. "improved" should be "improve". Page 5-4, line 25. I don't think young monkeys are called "pups". Page 5-6. Footnote "a" is not found in the table. Page B-5, Line 3. PPK should be PBPK. Page B-6. Table is hard to read (biochemical constants squished). Page B-7. Footnote g was not found in the table. Footnote k was not consistent with the text (i.e., source in table says Ernstgard et al., 2005, but Sedivec et al. (1981) in footnote). Page B-10, Line 26. The text about KMASC is not consistent with Table B-1.

	Page B-11. Listing one data source within the figure caption and one below the figure caption is confusing. Page B-13, Fig B-5. Legend says GD8, caption says GD9. Text (p B-14) says GD9. Text mentions 15 hr data, but none is evident in the figure. Page B-20, line 6. "know" should be "known". Page B-27, lines 1 and 2. This sentence needs a verb. Page B-27 line 6. "serious" should be "series". Page B-28. Daily dose should be in figure caption. Page B-35, lines 9 and 10, figure caption. Exponents should be properly formatted. Remove extra period in source citation. Page B-41, line 33. "scaled" should be "scale".

D2. Please identify any additional studies that would make a significant impact on the conclusions of the Toxicological Review and should be considered in the assessment of the non-cancer health effects of methanol.

Reviewer	Comments
Burbacher	Other than the new studies conducted by Dr. Wells' research group that were provided in an addendum, I do not know of additional studies on the reproductive and developmental toxicology of methanol.
Byczkowski	It seems that all important studies known to this reviewer (and several unimportant too) have been included in this Toxicological Review document.
Dorman	• The Toxicological Review has incompletely considered the rabbit studies published by Sweeting and coworkers. These studies were discussed in an amended document provided to this reviewer well after the June 17 draft was provided. These studies were not considered in the EPA's consideration of inter-species differences (i.e., are rat or mice studies appropriate). AS noted earlier, there is a lack of inclusion and exclusion criteria. For example, several studies that could be supportive are lacking from the document – e.g., Bolon et al (1993, 1994) There are also other studies, including work in monkeys, with aspartame that may be supportive (e.g., Reynolds). Since Table 3-2 includes results from aspartame exposure this does not seem to be a clear exclusion criterion. Likewise, search terms and databases examined have been poorly defined. • As noted by EPA the kinetics of methanol is heavily influenced by ventilation arte (page 3-18 and elsewhere). Leavens et al (2006) published data showing longitudinal changes in respiratory rate in pregnant rats that could be used with a gestational PBPK model.

	• Dorman and coworkers also measured deciduas methanol concentrations. This could be considered in the discussion that occurs on page 3-52. • The ethanol teratology literature has been largely ignored despite seom similarities in teratogenic response. This larger literature may help inform the MOA discussions in the draft document and help guide whether formaldehyde should be considered as the proximate teratogen.
McMartin	In the Addendum to the Draft Peer Review, there is discussion of interesting new results by Wells and coworkers (2010 and 2011). These investigations have compared the metabolism and pharmacokinetics of methanol in mice, rabbits and monkeys. As would be expected from numerous previous studies, monkeys showed a markedly higher level of formate accumulation and a slower rate of methanol elimination compared to mice. The interesting observation is that the rabbit appeared to be somewhere in the middle. As seen in Figure 4 of the Sweeting 2010 paper, methanol is cleared more slowly in the rabbit such that blood levels remain elevated for nearly 40 h (compared to only 20 h in mice). Similarly, formate blood levels are elevated for 40 h in rabbits and only 20 h in mice. The authors interpreted these data to indicate that methanol kinetics in rabbits more closely approximates those in primates (hence humans), so rabbits might be a better animal model than mice (rodents) for developmental studies. However, it is important to note that the peak formate level in the rabbit is no higher than that in mice, just the duration – the duration of formate elevation is strictly controlled by the elimination of methanol (once the latter is gone, formate is gone also). In that sense the rabbit is still different from the primate in terms of formate accumulation. However, formate doesn't appear to play a role in the developmental effects of methanol. As such, the slower elimination of methanol in the rabbit would make it a preferred model over mice for developmental studies – however, since the key study used for this RfC/RfD analysis is the NEDO rat study, the key comparison would be between rabbits and Sprague-Dawley rats (which also happen to eliminate methanol more slowly than do mice). The Discussion of these studies as well as the other ones by Wells and coworkers (esp the publication regarding the role of ROS in mediating the effects of methanol) needs to be improved in the document. Although the "results" of these studies are presented in the Addendum, the ramifications of these studies is not presented well or at all in the sections related to choice of POD, critical effect, etx.
Roberts	A description of additional studies was added as an Addendum to the Toxicological Review provided to the Panel. To these studies, the paper by Miller and Wells (Toxicol. Appl. Pharmacol. 252:55-61, 2011) should be added. This paper is germane to the discussion of potential mode of action for methanol developmental effects and also extrapolation of observations in mice to other species.
Salmon	No additional studies beyond those presented in the report (and the recent addendum) were identified.
Sweeney	Haffner HT, Wehner HD, Scheytt KD, Besserer K. The elimination kinetics of methanol and the influence of ethanol. Int J Legal Med. 1992; 105(2):111-4. This study presents blood concentration time course data for methanol infused iv in a healthy male volunteer. Data for 3 other individuals were not shown, but indicated similar blood half lives.

	Schmutte P, Bilzer N, Penners BM. Zur Nűchternkinetik der Beglietalkohole Methanol und Propanol-1. [Kinetics of the congeners methanol and propanol-1 in the absence of ethanol]. Blutalkohol. 1988; 25(3):137-42. This paper (in German) presents the blood methanol time course in 5 or 7 volunteers who ingested a methanol/drinking water mixture over a period of 15 minutes. Incorporation of human kinetic studies by the iv and oral routes in PBPK model development has the potential to change the estimates of the chemical-specific parameters in the PBPK model, and thus the HEC, HED, RfC and RfD.

D3. Please discuss research likely to substantially increase confidence in the database for *future* assessments of methanol.

Reviewer	Comments
Burbacher	It would be good for the current data to serve as the basis for future studies of methanol developmental toxicity. Future studies using different animal models from rodents to primates should focus on outcomes related to reproductive function, early sensorimotor development and object memory as well as changes in brain architecture and size.
Byczkowski	This reviewer is not aware of such studies.
Dorman	Development of a PBPK model that considers gestation and lactational exposure. There is a need to replicate the findings of the critical study used by NEDO including the inclusion of additional neuropathological and neurobehavioral assessments. Although additional monkey studies could be considered the Burbacher study is extremely robust and should receive more attention by EPA. Additional MOA data – especially studies designed to resolve whether formaldehyde is involved in the developmental effects seen following peri-natal methanol exposure are warranted. These studies should include the use of dual labeled material to confirm fetal exposure. Completion of surveys to examine blood methanol concentrations in the US population. Improved understanding of the mode of action of methanol and it's metabolites (formaldehyde and formate) in the teratogenic response seen in animals.
McMartin	Ideally, a study that fully characterizes methanol metabolism in the intact fetus and the dam using the rat as model would allow for a direct assessment of the role of fetal ADH and/or catalase in producing HCHO in the fetus (as opposed to the existing studies that only assess protein levels or activities using ethanol as substrate). Such a study would impact on the MOA for the developmental effects of methanol. Studies of the role of ADH and catalase in the metabolism of methanol by F-344 and Sprague-Dawley rats would clarify why there might be two saturable pathways in one strain but only one in the other (as implied by the PBPK model). These studies should also be done in a way to estimate the Km and Vmax for methanol metabolism to further refine the PBPK model with defined rather than fitted parameters.

	There is some concern that the NEDO rat studies are the only one to have reported decrease brain weight as a developmental effect of methanol. There are other studies showing various other effects of similar exposures to methanol, but none report decreased brain weight, nor any really dramatic CNS pathology (such as leading to real behavioral changes – the changes in several studies have been incredibly minor). Thus, it would be ideal to have an independent study to essentially repeat the NEDO-type exposures (i.e. 20+ hours per day instead of the 2.5 – 7 hours) to test if there are effects of methanol on the CNS under those circumstances (both on brain weight and on brain function/pathology). Again ideally it would be good to have a well conducted oral developmental study of methanol in order to produce sufficient data for an RfD (instead of having to do extrapolations from an inhalation study).
Roberts	Existing research clearly indicates the potential for methanol to produce developmental effects, but recent studies have suggested substantial differences among species and even strains, both in terms of susceptibility and type of effect. This creates enormous uncertainty in the extrapolation of effects from one species/strain to humans. Research to explain the basis for differences in species/strain developmental effects is essential. Also, sound dosimetry is compromised by lack of understanding of the proximate toxicant and mode of action for developmental toxicity. This must be resolved in order to more effectively utilize PBPK modeling for extrapolation across species.
Salmon	Although it is unlikely that additional extensive primate studies either could or should be undertaken, it is unfortunate that the main developmental and neurotoxicity studies in non-human primates are deficient in design, for instance using wild-caught animals and lacking a proper control group or dose-response assessment. It would be interesting to see further studies to illuminate the relative sensitivity of rodents and primates to chronic methanol toxicity, especially with regard to developmental and neurotoxicity endpoints. Possibly some further studies *in vitro* would be illuminating. The lack of an assessment of neurobehavioral impacts on rodent development is a significant data gap, as noted by the Agency in their selection of a database uncertainty factor. Given the reliance of the RfC on an anatomical measure of developmental neurotoxicity (brain weight reduction) which might prove less sensitive than a functional evaluation, this is a significant deficiency which does not seem too difficult to remedy.
Sweeney	Inhalation kinetic data for Sprague-Dawley rats appear to be limited. Monkey studies with longer exposure durations and similar endpoints could be informative. Additional mode-of-action motivated studies would be helpful.

Bonus Charge Question: Please comment on the proposed RfD and RfC values for their intended use in risk assessment. Are these numbers more conservative than they need to be to protect public health?

Reviewer	Comments
Burbacher	[no comments provided]
Byczkowski	The RfD developed in the present document is numerically almost the same as the previous one, already listed in IRIS data base. Although, the upper bound on concentration of normal physiological "background" of methanol in humans has not been determined in this document, it seems that the exposure to methanol at the proposed RfD or RfC level may produce internal concentration not much different from the physiological background. This makes both, the existing and the proposed reference toxicity values for methanol very conservative. While the overall goal of developing reference toxicity values is to protect public health, perhaps, the revision of the current document would give an opportunity to U.S. EPA to derive RfC and RfD that would be not only health-protective, but also realistically close to the no-adverse-effect level in humans, with reasonable margin of safety and appropriate confidence.
Dorman	As defined by the EPA, the RfC is "An estimate of a continuous inhalation exposure for a given duration to the human population (including susceptible subgroups) that is likely to be without an appreciable risk of adverse health effects over a lifetime." Practically speaking some individuals also perceive the RfC as a point of demarcation, with exposures in excess of the RfC representing an exposure where adverse health effects *may occur*. Thus, although an RfC = 0 meets the first working definition, this concentration ("0") is overly conservative. My greatest reservation about the EPA's proposed RfC value relates to the observation that exposures at (or slightly above) the RfC would result in a change in blood methanol concentration that falls within the range of normal values seen in people. This value becomes difficult to defend scientifically for an endogenous chemical like methanol. The proposed RfC value is the result of systematic application of multiple conservative estimates (often lacking transparency or scientific justification) and seemingly arbitrary application of additional default uncertainty factors. Some of these rather conservative steps included: • Use of a SD versus 5% BMDL value [~2X] • Use of a single 6 hr data point rather than consider looking at an estimate based on all time points in the NEDO study • Inclusion of a modifying factor for database uncertainties This led to the derivation of an RfC that is overly conservative. Ultimately, the EPA team should ask the question when a change in blood methanol concentration may lead to a toxicologically significant effect. This approach was considered by the NTP CERHR group that considered a blood methanol concentration of < 10 mg/dl to not be associated with adverse developmental effects. This determination considered the available toxicity data from people and animals and normal blood methanol concentrations seen in people with various dietary inputs.

McMartin	The RfD and RfC values have been appropriately derived based on the BMD/PBPK analysis utilizing "standard EPA procedures", but the resulting values lack scientific credence and are not logical in the sense of the exposures expected for humans. The problem scientifically is that the resulting values reflect a methanol exposure that does not increase the blood methanol concentration in the exposed human. Because of the background level of methanol in all humans lies in the range of 2 mg/L, the projected increase in methanol level from the RfC/RfD exposure is only 0.04 mg/L, i.e. a level that is really indistinguishable from the background. The implications of this include that all humans would be susceptible to developmental effects of methanol no matter what exposure they had experienced – not a suitable endpoint for risk assessment. As stated above, presuming that endogenous levels of methanol do not contribute to adverse effects and an exposure does not produce an increase above background levels, how can that exposure lead to an adverse effect? The conundrum occurs because the PBPK model itself has built-in conservatism, the BMD calculation has built-in conservatism and then a 100-fold uncertainty is applied. All of these factors contribute to bring the "RfC/RfD exposure" down to the levels where there is essentially no exposure-induced increase in methanol levels above the endogenous, background level, which means there is essentially no risk. So in this case of an endogenous chemical, the numbers are more conservative than necessary.
Roberts	The process of developing these RfC and RfD values has produced a result that is counter-intuitive, implying that individuals with no unusual methanol exposure may be at risk of developmental effects. That's implausible, and clearly signals the need to re-evaluate how to consider background methanol concentrations in the development of credible toxicity values. What is missing from this and other IRIS toxicological reviews is an assessment, after going through the process of RfC and RfD development, whether the resulting values make sense, i.e., are they logical in the context of exposures and effects anticipated in human populations? For many chemicals, this may be impossible to judge; there simply is not enough information about exposure or exposed populations to make a determination. For more familiar chemicals, if there is some question about the validity of the toxicity values, it should incumbent upon the toxicological review to make a case why the toxicity values are reasonable. Public comments have argued that the proposed methanol RfC and RfD values correspond to doses commonly experienced by the public (e.g., one small glass of orange juice daily), and are projected to produce increases in blood methanol concentrations for most individuals that would keep them with the background range. The NTP-CERHR methanol panel considered common exposures to methanol and concluded that they pose no immediate concern for developmental toxicity (NTP-CERHR, 2003). This appears to be a reasonable conclusion and creates a real credibility problem for the proposed methanol RfC and RfD.
Salmon	The issue about whether these values are too conservative arises because of the observation that, using the PBPK model to predict the additional levels of blood methanol levels at the RfC or RfD, it appears that these levels result in a relatively small (1 to 15%) increment in blood methanol concentrations relative to the background levels resulting from metabolism and/or dietary exposures. A point that needs to be emphasized is that the RfC and RfD are specifically defined as levels at which the risk assessor can be reasonably confident that adverse effects will not appear. They are not threshold levels at which effects might start to appear, and there is

	no generally accepted method, other than the pragmatic consideration of hazard index values, for determining what risks may exist from exposures above the RfC or RfD. However, it is certainly true that this sort of calculation may affect the popular perception of the reasonableness of the calculation, and need to be addressed in discussion of the proposed values. The proper approach to the question about "normal" background values relative to the RfC and RfD is to ask whether the observed background levels place any constraints on the values of uncertainty factors used in derivation of the protective levels. If a given background level is seen as "safe" (or even beneficial in the case of agents such as fluoride or essential metals), it obviously does not make sense to set values of UF_Aor UF_H which cause the protective levels substantially lower than this level, since this background level can be seen as providing additional evidence on human sensitivity to the agent in question. Unfortunately in the case of methanol this is complicated by the fact that observed background levels in humans appear to show both substantial inter-individual variation and substantial temporal variation in a single individual. Ordinarily, a "safe" level might be identified by considering a point somewhere near the high end of the distribution of background levels. This might be sustainable if you believe that the adverse developmental effects of methanol are strictly caused by methanol itself. However, in view of the uncertainties as to fetal metabolism, mode of action and contribution of diet and individual metabolic or toxicodynamic differences which are identified in the report it seems very unwise to conclude that high-end exposures which are apparently safe for some individuals are *necessarily* safe for all. The assumption that current background levels of methanol are without effect seems plausible, but it is not clear that there have been any analytical investigations of this issue. At the very least an uncertainty factor is needed to reflect these concerns, which therefore indicates that the proposed values for RfC and RfD are not necessarily unreasonable. The Agency needs to address this discussion, which is not covered in the draft report.
Sweeney	In a word: Yes. It does not seem logical to conclude that methanol exposures that increase human blood concentrations from ~1.8 mg/L to ~1.84 mg/L constitute a threshold for meaningful increases human risk. (See A.2.) How did we get here? The answer seems to be: the combination of excessive UFs with a point of departure that is lower than it should have been (due to an incomplete consideration of the rat inhalation toxicokinetic database). The proposed UFs are excessive in light of the available toxicological database and minimal contribution of additional methanol to the body burden at concentrations/doses relevant to the RfC and RfD (see B4). More appropriate choices (higher point of departure, lower UFs) may yield RfC and RfD values that are less of a departure from common sense, yet still provide adequate health protection. The BMDLs were derived on the basis that changes in brain weight should not exceed 1 standard deviation of the natural background variation. It is not unreasonable to extend that reasoning to increases in a naturally occurring background chemical measured in blood to provide a reality check on the reference values derived by the POD/UF approach.

Reviewer SpecificComments

Additional Comments Submitted by Dr. Thomas Burbacher

Comments on the Burbacher et al monkey study

The document's treatment of the Burbacher et al monkey study seems inconsistent and in some cases is inaccurate. The statements below should be reviewed and edited for clarity.

Page 3-13 –It is not clear why the Burbacher et al study should be criticized for not being "relevant to persons who are folate deficient". This could be said for most of the studies reviewed in the document. This statement needs to be clarified or deleted.

Page 3-17 –The references provided for the lack of consistent effects observed in "primate exposure studies" are mostly not primate studies. The long list of references gives the impression that there are a lot of studies in primates when in fact only 1 reference provided is a primate study.

Page 3-24 –The document includes the Burbacher et al study on the list of studies that provides PK data for methanol (3-25). Subsequent sections of the document include the study (3-46) and at other times exclude the study when discussing PK issues (3-29, 3-42, 3-51 summary and conclusions). The current document should be reviewed to provide a more consistent discussion of the contribution of this study to methanol PK issues.

Page 4-36 –The document states that the decreased length of pregnancy observed in the Burbacher et al study "was largely due to complications of pregnancy requiring Cesarean section deliveries". This is not accurate. The decreased length of pregnancy is observed in only vaginally delivered animals. The authors should review this issue and adjust their comments. NOTE: This statement is repeated many times in the document (examples 4-37, 5-4, 5-5, 5-6, 5-27) and in several tables (4-79, 5-6). The document also points out that the study includes a mixture of "feral-born and colony-bred animals". It is not clear what point is being made here. Is there a problem with inclusion of feral born animals? If so, this would be setting a new standard for these studies since this is the norm. In addition, the study counter-balanced the adult females on age, weight, and origin and there were no effects that were dependent on origin. This statement should be clarified or deleted.

Page 4-55 –It is not clear how the decision was made to use unadjusted VDR responses for females only. The unadjusted data for males provides a better dose-response relationship. This should be clarified.

Page 4-56 –The primary analysis for the recognition memory study is focused on whether or not the various groups exhibit a novelty preference. The lack of a novelty preference in the exposed groups is consistent with previous data from this task with other high risk infant groups (ethanol exposed, methylmercury exposed, premature). I would recommend that the authors consider including a behavioralist in the review of the Burbacher study, particularly one familiar with human developmental studies.

Page 4-82 –The document indicates that the Burbacher study is provides evidence for species-to-species variation in susceptibility. The synopsis of the study then goes on to repeat many of the issues discussed. This discussion should be reviewed to address the concerns stated above.

Additional Comments Submitted by Dr. Andrew G. Salmon

EDITORIAL COMMENTS

Page 3-7: The new version of Figure 3-1 in the addendum is a great improvement on the previous figure, which was confusing and uninformative.

Some minor typographic and editorial anomalies were noted while reading the document. These should be rectified in the final version:

Page 3-12: Footnote 9, final sentence: "… this too was not statistically significant." What other non-significant result is referred to here? Clarify, or remove the word "too".

Page 4-4 line 32, page 4-5 line 2, page 4-7 line 1. The antidote/drug fomepizole is referenced on these three occasions, but it is not until page 4-41 line 2 that it is explained that this is chemically 4-methylpyrazole, and that it is an inhibitor of ADH1. It would be useful to have this definition of the chemical name and mode of action pointed out at the first reference rather than the last.

Pages 3-20 to 3-22. The description of the model by Ward et al. (1997) appears twice (as Sections 3.4.2.1 and 3.4.2.3), both appearances being identical except for Table 3-8 which appears only in the second instance. Delete the redundant entry and re-number the sections.

Page 4-1, line 22 refers to "lentiform nuclei". Are these the same as the "lenticular nuclei" referenced on Page 4-2 line 5? If so the terminology should be consistent: if not the difference should be explained.

Page 4-2 lines 18-19. "…Benton and Calhoun (1952) reported on methanol's visual disturbances". What did they say? Also this sentence could be phrased better – it was the 320 individuals who had the visual disturbances, not the methanol.

Page 4-18, line 27. What are "proximal uniferous tubules" in the kidney? Is this a typo for uriniferous? Most people just call them proximal tubules.

Pages 4-30 to 4-31. Table 4-2 would be much easier to read if it did not split over two pages. Some slight abbreviation might be necessary to achieve this.

Page 4-77 line 7. "The data are summarized separate sections…" Need the word "in" after "summarized".

Page B-41 line 33. "did not scaled" should read "scale".

Appendix A: Individual Reviewer Comments

COMMENTS SUBMITTED BY

Thomas M. Burbacher, Ph.D.

Peer Review Meeting of EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Comments Submitted by Dr. Thomas M. Burbacher

(A) Toxicokinetics and PBPK Modeling

A PBPK model developed by EPA based on models by Ward et al. (1997) and Fisher et al. (2000) was utilized in the Toxicological Review of Methanol.This model is described in Section 3 and a detailed description of the EPA model modifications, evaluation, and application are found in Appendix B.The PBPK model modified by EPA can estimate internal dose levels due to exogenous methanol exposure (i.e., doses above background).This modified methanol PBPK model was first applied to predict internal doses in experimental animals under bioassay conditions.Benchmark dose (BMD) modeling, using internal doses as exposure metrics, was then used to identify internal-dose points of departure (PODs) from the animal data.Finally the human PBPK model was used to identify human equivalent concentrations (HECs) or doses (HEDs) for each internal-dose POD.

Note: Background methanol levels have been subtracted by study authors from most of the mouse and rat pharmacokinetic data and those background levels are not reported. Since the goal is to predict risk above background, the EPA subtracted background levels from the pharmacokinetic data where it was otherwise included, to obtain a consistent total data-set for use in developing the PBPK models.The underlying assumption is that non-cancer risks from methanol exposure are due to increases in the levels of methanol or its metabolites above background.

A1. Please comment on the scientific soundness of the PBPK model used in this assessment.

The assessment includes detailed reviews of PBPK models for mouse, rat, and humans. The reviews are comprehensive but it is not clear to this reviewer what the strengths and weaknesses are for each model and why the nonhuman primate model was not included in the final model development. Some clarification of the process for evaluating the usefulness of each model for the assessment and why the nonhuman primate model was not included would be helpful.

A2. Please comment on the scientific justification for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks.

The rationale for this approach is weak. To be consistent with longstanding assumptions used by EPA and others, the critical factor related to toxicity is the total level of methanol from all sources, or cumulative exposure. The rational for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks is not clearly stated and it gives the impression that cumulative exposures from different sources are not important. On page 3-28 it states that the modeling that includes background levels was estimated to have "minimal impact" on the dose extrapolations. It would seem that if this is the case, these "more complex" models would also be more rigorous and appropriate for use in this assessment.

A3. The PBPK modeling effort assumed similar methanol pharmacokinetics between pregnant and non-pregnant animals. Please comment on the adequacy of the dose-metric extrapolation based on a PBPK model for non-pregnant adults (i.e., no fetal compartment) for predicting risks associated with fetal/neonatal brain concentrations of methanol.

The review of data related to methanol pharmacokinetics between pregnant and non-pregnant animals provides sufficient evidence for the assumption of similarity. There are data in nonhuman primates that indicate little change during pregnancy.

A4. EPA assumes limited methanol metabolism in the fetus because of limited alcohol dehydrogenase (ADH) activity in the human fetus, limited catalase and ADH activity in fetal rodents, and existing pharmacokinetic data that show nearly equal concentrations in maternal blood vs. the fetal compartment.Please comment on the validity of this assumption given the lack of data regarding potential alternate metabolic pathways in the fetus.

This seems like a reasonable assumption given the limited data available.

A5. Please comment on the scientific justification of the extrapolation approach from rats to humans for in-utero and neonatal lactational and inhalation exposures.

As I mentioned above, the reviews of PBPK models for mouse, rat, and humans are comprehensive but it is not clear to this reviewer what the strengths and weaknesses are for each model and why the nonhuman primate model was not included in the final model development. Some clarification of the process for evaluating the usefulness of each model for the assessment and why the nonhuman primate model was not included would be helpful. There are also issues related to using the NEDO studies which included neonatal exposures that continue to be problematic, given the lack of data on lactational and early postnatal inhalation exposure to methanol.

(B) Inhalation Reference Concentration (RfC) for Methanol

B1. A chronic RfC for methanol has been derived from a perinatal inhalation study of the effects from exposing rat dams and pups to methanol during gestation and lactation (NEDO, 1987).Reference values from mouse (Rogers et al., 1993) and monkey (Burbacher et al., 1999; 2004) developmental studies, were also derived and discussed, but were not chosen for the RfC. Please comment on whether the selection of the principal study has been scientifically justified.

After reviewing the previous panel comments of the NEDO study, I do not believe that this is the most appropriate study for derivation of the RfC. The review indicated that there continue to be questions regarding the procedures used in the NEDO study (in utero and postnatal exposures, litter effects, etc) that make it difficult to evaluate the study for RfC derivation. The discussion on page 5-10 regarding the complications that arise from using the NEDO study where exposure was both gestational and postnatal postulates a number of assumptions that are supported by little or no data. Data on lactational transfer and early postnatal inhalation exposures are limited. A note: On page 5-15 the document indicates that the monkey VDR effects are an example of prenatal and continuing postnatal exposure effects. This is not accurate since the monkeys were only exposed prenatally.

The study by Rogers et al., 1993 would seem to be the most appropriate choice at this time. The study is scientifically sound and robust. Exposures are limited to the prenatal period and the outcomes are clear. While I do not agree with several comments regarding the Burbacher et al monkey study, 1999, 2004 (see comments below), I do agree that the monkey study should not be used for derivation of the RfC due to the lack of a dose-response function for the major effects.

B2. Reduction of brain weight at 6 weeks postnatally as reported in the NEDO (1987) developmental rat study was selected as the critical effect.Please comment on whether the rationale for the selection of this critical effect has been scientifically justified.Please identify and provide the rationale for any other endpoints (e.g., other reproductive and developmental effects reported in mouse and monkey studies) that should be consideredin the selection of the critical effect.

As mentioned above, the Rogers at al. 1993 study would be a more appropriate study to use for deriving the RfC, using increased incidence of cervical ribs as the critical effect.

B3. Benchmark dose modeling of decreased pup brain weight relative to maternal internal methanol doses predicted by the PBPK model was used to derive the point of departure (POD) for the RfC.Has the BMD/PBPK approach been appropriately conducted?Has adequate justification been provided for the selected internal dose metric, i.e., area under the curve (AUC) for methanol, in the blood of dams?Please identify and provide the rationale for any alternative approaches for the determination of the POD, including choice of another dose metric (e.g., methanol metabolized), and discuss whether such approaches are preferred to EPA's approach.

This is not my area of expertise so I will not comment.

B4. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the POD for the derivation of the RfC.It is assumed that these UFs account for variability in methanol dosimetry among human newborns following gestational and lactational exposure, and for uncertainty regarding the ratio of newborn-dose to maternal-dose in humans. Please comment on these assumptions and on the scientific justification for the selected UFs.

The rationale for selecting the uncertainty factors is consistent with EPA policy.

(C) Oral Reference Dose (RfD) for Methanol

C1. EPA concluded that the oral RfD should be derived using a route-to-route extrapolation from the more extensive inhalation database given the paucity of oral toxicity data.Please comment on whether the rationale for this approach has been scientifically justified and clearly explained. Please identify and provide the rationale for any alternative approaches for the determining the RfD and discuss whether such approaches are preferred to EPA's approach.

This is not my area of expertise so I will not comment.

C2. A PBPK model was used to derive the RfD via a route-to-route extrapolation, in which the internal-dose POD used for the derivation of the RfC based on data from the NEDO (1987) study was extrapolated to human oral exposure levels using the human PBPK model.Please comment on whether the rationale for this approach has been scientifically justified.Has adequate justification been provided for the selected internal dose metric, i.e., AUC for methanol, in the blood of dams?Is the PBPK model suitable for extrapolation of fetal and neonatal endpoints to human oral exposures?Please provide a detailed explanation.

This is not my area of expertise so I will not comment.

C3. EPA applied the same UFs to the POD for the derivation of the RfD as for the RfC.Please comment on the rationale for the selection of the UFs.

Again, the rationale for selecting the uncertainty factors is consistent with EPA policy. Although I would assume there would be more uncertainty due to the lack of oral exposure data.

(D) General Charge Questions

D1. Is the Toxicological Review logical, clear and concise?Has EPA clearly synthesized the scientific evidence for non-cancer hazards?

There is probably not much that can be done about it since I assume there is a format for these reports, but I found the assessment to be very repetitive. I would like to commend the authors of the assessment for a very comprehensive review. The only comment I have aside from the repetitive nature of the review, is that it may be good to provide some discussion and/or figure related to “decision trees” for major decisions that were made in terms of choices of approaches, studies, outcomes, etc. to make these choices more transparent than they currently are. The comments regarding the Burbacher et al study seemed inconsistent and at time inaccurate (see below for specific comments regarding the Burbacher et al study).

D2. Please identify any additional studies that would make a significant impact on the conclusions of the Toxicological Review and should be considered in the assessment of the non-cancer health effects of methanol.

Other than the new studies conducted by Dr. Wells’ research group that were provided in an addendum, I do not know of additional studies on the reproductive and developmental toxicology of methanol.

D3. Please discuss research likely to substantially increase confidence in the database for *future* assessments of methanol.

It would be good for the current data to serve as the basis for future studies of methanol developmental toxicity. Future studies using different animal models from rodents to primates should focus on outcomes related to reproductive function, early sensorimotor development and object memory as well as changes in brain architecture and size.

Comments on the Burbacher et al monkey study

The document's treatment of the Burbacher et al monkey study seems inconsistent and in some cases is inaccurate. The statements below should be reviewed and edited for clarity.

Page 3-13 –It is not clear why the Burbacher et al study should be criticized for not being "relevant to persons who are folate deficient". This could be said for most of the studies reviewed in the document. This statement needs to be clarified or deleted.

Page 3-17 –The references provided for the lack of consistent effects observed in "primate exposure studies" are mostly not primate studies. The long list of references gives the impression that there are a lot of studies in primates when in fact only 1 reference provided is a primate study.

Page 3-24 –The document includes the Burbacher et al study on the list of studies that provides PK data for methanol (3-25). Subsequent sections of the document include the study (3-46) and at other times exclude the study when discussing PK issues (3-29, 3-42, 3-51 summary and conclusions). The current document should be reviewed to provide a more consistent discussion of the contribution of this study to methanol PK issues.

Page 4-36 –The document states that the decreased length of pregnancy observed in the Burbacher et al study "was largely due to complications of pregnancy requiring Cesarean section deliveries". This is not accurate. The decreased length of pregnancy is observed in only vaginally delivered animals. The authors should review this issue and adjust their comments. NOTE: This statement is repeated many times in the document (examples 4-37, 5-4, 5-5, 5-6, 5-27) and in several tables (4-79, 5-6). The document also points out that the study includes a mixture of "feral-born and colony-bred animals". It is not clear what point is being made here. Is there a problem with inclusion of feral born animals? If so, this would be setting a new standard for these studies since this is the norm. In addition, the study counter-balanced the adult females on age, weight, and origin and there were no effects that were dependent on origin. This statement should be clarified or deleted.

Page 4-55 –It is not clear how the decision was made to use unadjusted VDR responses for females only. The unadjusted data for males provides a better dose-response relationship. This should be clarified.

Page 4-56 –The primary analysis for the recognition memory study is focused on whether or not the various groups exhibit a novelty preference. The lack of a novelty preference in the exposed groups is consistent with previous data from this task with other high risk infant groups (ethanol exposed, methylmercury exposed, premature). I would recommend that the authors consider including a behavioralist in the review of the Burbacher study, particularly one familiar with human developmental studies.

Page 4-82 –The document indicates that the Burbacher study is provides evidence for species-to-species variation in susceptibility. The synopsis of the study then goes on to repeat many of the issues discussed. This discussion should be reviewed to address the concerns stated above.

COMMENTS SUBMITTED BY

Janusz Z. Byczkowski, Ph.D., DABT

Peer Review Meeting of EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Comments Submitted by Dr. Janusz Z. Byczkowski

(A) Toxicokinetics and PBPK Modeling

A1. Please comment on the scientific soundness of the PBPK model used in this assessment.

By definition, science compiles, completes and systematically organizes knowledge, providing as detailed description of the investigated phenomena as currently possible. In contrast, science-based human health assessments and toxicological reviews aim at supporting the U.S. EPA regulatory activities designed to protect public health. Thus, the models used in risk assessment serve only a relatively narrow goal of estimating the health protective exposure levels of xenobiotics which are most likely without adverse effect to humans. To fulfill this goal, the modelers often have to choose shortcuts, simplifications, surrogate dose metrices, and to make health-protective assumptions for PBPK models, which may be unacceptable for scientific dissertation, but if technically correct - according to the rule of parsimony, they can be applied as tools in approximation of dose and/or response to xenobiotics for EPA regulatory purpose.

The PBPK model used in this assessment includes at least three major shortcuts: i) it is lacking a detailed quantitative description/distribution of endogenous levels of methanol (background); ii) lacking qualitative and quantitative description of metabolites; and iii) lacking quantitative description of lactational transfer of methanol and its concentration in the postnatal brain. While such omissions could be perceived as a deficiency in scientific description of the phamacokinetics of methanol, this PBPK model seems to be adequate for the risk assessment purpose, for which it was developed. Obviously, this and any other PBPK model is only as good, as good were the data used for its validation.

As stated in the reviewed document on P. 3-24 (L# 5 - 7): "...it was determined that a modified Ward et al. (1997) model, with the addition of a lung compartment as described by Fisher et al. (2000), should be used for the purposes of this assessment..." The choice and combination of these two PBPK models seems to be optimal. Even though, the combined "hybrid" model has been significantly simplified, it is still adequate for use in dose-effect modeling and interspecies extrapolations. The ACSL codes of the model have been listed in the Appendix B, including *.CSL and *.CMD files (and even the runtime *.m files). The PBPK model seems to be appropriately constructed, using the principle of parsimony, and it is very well documented. While the use of "drinking tables" (P. B-49; L#39 to P. B-50; L#12) cannot be considered to be "mathematically elegant", it represents a pragmatic solution to the problem of modeling different patterns of drinking/feeding by different species.

A2. Please comment on the scientific justification for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks.

The assumption that adverse effects of methanol exposure appears at internal concentrations higher than its physiological background is correct, analogously to many other chemicals - essential at physiological

concentrations but deleterious at high external doses. The physiological levels of one carbon groups are metabolically necessary for the organism and they do not produce adverse effects. Since the U.S EPA can regulate only the external exposures - not the normal endogenous concentrations of chemicals, apparently a reasonable decision was made to subtract the background level from data used in the quantification of pharmacokinetics of methanol. On the other hand, any exposure to external methanol, adds up to the existing background, increasing concentration of methanol in the target tissue. So, the upper bound on background concentrations of methanol in target tissue should be carefully evaluated and used consequently. The lack of determination of the upper statistical bound on normal physiological concentrations of methanol in relevant species, including humans, can be considered to be a major deficiency of the reviewed document.

The justification for selecting a no-background model, over the PBPK model that does include background, has been provided in the document (Section 3.4.3.2.1 and P. 3-28; L# 4 to 7): "*...more complex PBPK modeling required to include background levels was estimated to have a minimal impact on dose extrapolations, the use of simpler methanol models that do not incorporate background levels is considered adequate for the purposes of this assessment*..."

Codes for the background level, blocked in the final simulations, have been incorporated in the PBPK model, as documented on P. B-48; L#9 to 55 and P. B-49; L# 1 to 3; and as stated on P. 6-3; L# 21 to 24: "*...This assessment focuses on the determination of noncancer risk associated with exogenous methanol exposures that increase the body burden of methanol or its metabolites (e.g., formate, formaldehyde) above endogenous background levels*..." However, in the simulations whose results are listed in the Table B-5, a background level of 2 mg/L has been set to model human internal concentration from inhalation (P. B-92; L# 29) but not from the oral exposure (P. B-92; L# 55).

A3. The PBPK modeling effort assumed similar methanol pharmacokinetics between pregnant and non-pregnant animals. Please comment on the adequacy of the dose-metric extrapolation based on a PBPK model for non-pregnant adults (i.e., no fetal compartment) for predicting risks associated with fetal/neonatal brain concentrations of methanol.

Due to its physicochemical properties, "*...methanol penetrates cellular membranes readily and distributes throughout total body water*..." (c.f. P. 3-11; L# 14). This was confirmed by the experimental data, which according to the statement on P. 3-10 L# 20 to 22: "*... as a whole suggested that the distribution of orally and i.v. administered methanol was similar in rats versus mice and in pregnant rodents versus NP rodents*..." Figure 3-3 (P. 3-11) provided further evidence that within the relevant range of concentrations, the ratio of methanol concentration in amniotic fluid to concentration of methanol in maternal blood is nearly linearly 1 to 1. Thus, the dose metric extrapolation, based on non-pregnant adults to predict internal concentrations of methanol in fetal/neonatal brain, seems to be adequate, as explained and substantiated in the document.

A4. EPA assumes limited methanol metabolism in the fetus because of limited alcohol dehydrogenase (ADH) activity in the human fetus, limited catalase and ADH activity in fetal rodents, and existing pharmacokinetic data that show nearly equal concentrations in maternal blood vs. the fetal compartment. Please comment on the validity of this assumption given the lack of data regarding potential alternate metabolic pathways in the fetus.

Assumption of limited methanol metabolism in the fetus is consistent with the demonstrated low fetal metabolic activity for several other xenobiotics similar to methanol. As explained in the document P. B-3; L# 5 to 10: "...*The fact that measured fetal blood levels are virtually identical to maternal levels for methanol (and ethanol) tells us that the rate of metabolism in the fetus is not sufficient to significantly reduce the fetal concentration versus maternal...*" So, even though the contribution of potential alternate metabolic pathways in the fetus remains uncertain, this reviewer agrees with the simplified assumption, as explained in the document (P. B-2; L#19 to 21): "...*Because the maternal blood:fetal blood partition coefficients were near 1, there was no need to explicitly model fetal kinetics; they will be equivalent to maternal blood kinetics...*" Obviously, this assumption holds only if the parent compound (methanol itself) is indeed responsible for deleterious effects in the fetal/postnatal brain. As far as this reviewer is aware, to date, no convincing study explaining potential teratogenic mechanism of action of methanol was published.

A5. Please comment on the scientific justification of the extrapolation approach from rats to humans for in-utero and neonatal lactational and inhalation exposures.

As explained in the section B.2.9 of the document (P. B-39, L#28 to P. B-41, L#2): "...Although the developmental endpoints of concern are effects which occur during in utero and (to a lesser extent) lactational exposure, it is not necessary for a MeOH PBPK model to specifically describe pregnancy (i.e., specify a fetal/gestational/conceptus compartment) and lactation in order for it to provide better cross-species extrapolation of risk than default methods..."

Due to its physicochemical properties, it was logical to postulate that the concentration ratio of methanol in maternal vs fetal compartments should be consistently close to the unity, across different species. This allowed the modelers to simplify further the PBPK model and to perform interspecies extrapolations under assumption that the maternal compartment realistically represents the fetal one. Again, this assumption holds only if the parent compound (methanol itself) is indeed responsible for deleterious effect in the fetal/postnatal brain.

The assumption of 1 to 1 ratio do not necessarily holds for lactational transfer of methanol, but it has been explained in the document (P.B-40, L# 18 to 36 and P. B-41, L# 1 to 2) that: "... *While lactational exposure is less direct than fetal exposure and blood or target-tissue levels in the breast-feeding infant or pup are likely to differ more from maternal levels, the health-effects data indicate that most of the effects of concern are due to fetal exposure, with only a small influence due to postbirth exposures...*"

Whereas it would be prudent to add the lactatonal transfer of methanol in to PBPK model, but as discussed in the answer to A1 (above), for the purpose of risk assessment this may be not necessary.

(B) Inhalation Reference Concentration (RfC) for Methanol

B1. Please comment on whether the selection of the principal study has been scientifically justified.

The selection of NEDO (1987) as the study principal in developing RfC, seems to be reasonable, even though the selection has been justified on practical/technical grounds rather than scientific (c.f. P.5-5, L# 10 to 24): "*...Taking into account the limitations of the studies available for quantification purposes, decreased brain weight at 6 weeks in male Sprague-Dawley rats exposed throughout gestation and the postnatal period (NEDO, 1987) was chosen as the critical effect for the purposes of this dose-response assessment as it can be reliably quantified and represents both a sensitive organ system and a key period of development*..."

B2. Please comment on whether the rationale for the selection of this critical effect has been scientifically justified. Please identify and provide the rationale for any other endpoints (e.g., other reproductive and developmental effects reported in mouse and monkey studies) that should be considered in the selection of the critical effect.

For justification of the critical effect and principal study, see the answer to B1 (above). This reviewer is not aware of any more appropriate end point.

B3. Has the BMD/PBPK approach been appropriately conducted? Has adequate justification been provided for the selected internal dose metric, i.e., area under the curve (AUC) for methanol, in the blood of dams? Please identify and provide the rationale for any alternative approaches for the determination of the POD, including choice of another dose metric (e.g., methanol metabolized), and discuss whether such approaches are preferred to EPA's approach.

The BMD/PBPK approach has been appropriately applied in the derivation of RfC, using health protective default (change of one S.D. from control mean) to determine BMR, according to the U.S. EPA guidance. The selection of AUC of methanol in maternal blood as a surrogate dose metric for dose-effect modeling of postnatal changes in brain also seems to be technically correct. Again, as discussed in the answer to A1 (above), the selection has been justified on practical/technical grounds. Without understanding of the exact mechanism of action of the chemical, selection of any surrogate dose metric is somehow speculative.

B4. Please comment on these assumptions and on the scientific justification for the selected UFs.

The following uncertainty factors, totaling UF of 100-fold, were applied in the development of RfC: UF_H = 10; UF_A = 3; and UF_D = 3. Even though, selection of uncertainty factors is usually arbitrary to some extent, the justification provided in the document seems to be convincing and in compliance with the health-protective U.S. EPA guideline.

(C) Oral Reference Dose (RfD) for Methanol

C1. Please identify and provide the rationale for any alternative approaches for the determining the RfD and discuss whether such approaches are preferred to EPA's approach.

The selection of route-to-route extrapolation from inhalation study to develop oral reference value is clearly explained in the document (e.g. on P.5-23 L# 11 to 15): "*...The limited data for oral administration indicate similar effects as reported via inhalation exposure (e.g., the brain and fetal skeletal system are targets of toxicity). Methanol has been shown to be rapidly and well-absorbed by both the oral and inhalation routes of exposure (CERHR, 2004; Kavet & Nauss, 1990). Once absorbed, methanol distributes rapidly to all organs and tissues according to water content, regardless of route of exposure...*"

This reviewer is not aware of any more appropriate approach and/or study that would be relevant to human chronic oral exposure to methanol. However, this is remarkable, that the RfD derived in this document is numerically almost the same as the previous one, derived from rat oral subchronic study.

C2. Please comment on whether the rationale for this approach has been scientifically justified. Has adequate justification been provided for the selected internal dose metric, i.e., AUC for methanol, in the blood of dams? Is the PBPK model suitable for extrapolation of fetal and neonatal endpoints to human oral exposures? Please provide a detailed explanation.

The reviewer's comments in answers to B1 through B3 (above) regarding RfC, are relevant also to this question about RfD. The rationale provided in the document for use of route-to-route extrapolation has been quoted in the answer to C1 (above). Regarding the AUC for methanol in blood of dams, selected as an internal dose metric, again - "*without understanding of the exact mechanism of action of the chemical, selection of any surrogate dose metric is somehow speculative*" (quoted from answer to B3, above).

C3. Please comment on the rationale for the selection of the UFs.

The health-protective UF totaling 100-fold, the same as in developing RfC, has been applied to RfD. As I stated in the answer to B4 (above): "*Even though, selection of uncertainty factors is usually arbitrary to some extent, the justification provided in the document seems to be convincing and in compliance with the health-protective U.S. EPA guideline*".

(D) General Charge Questions

D1. Is the Toxicological Review logical, clear and concise? Has EPA clearly synthesized the scientific evidence for non-cancer hazards?

The reviewed Toxicological Review document is logical and mostly clear. However definitely it is not concise. It provides not only the synthesis of evidence for noncancer health hazards, but it also attempts to describe most of the extensive research related to methanol toxicity, sometimes in a redundant manner. The length and the redundancy of the document are, perhaps, resulting from the formal structure required for IRIS Toxicological Reviews by U.S. EPA. It seems that in the revision of the current document, the U.S.

EPA could use the recent NAS recommendation that "*EPA should edit documents to reduce the text volume and address redundancies and inconsistencies*".

D2. Please identify any additional studies that would make a significant impact on the conclusions of the Toxicological Review and should be considered in the assessment of the non-cancer health effects of methanol.

It seems that all important studies known to this reviewer (and several unimportant too) have been included in this Toxicological Review document.

D3. Please discuss research likely to substantially increase confidence in the database for future assessments of methanol.

This reviewer is not aware of such studies.

Bonus question: Please comment on the proposed RfD and RfC values for their intended use in risk assessment. Are these numbers more conservative than they need to be to protect public health?

The RfD developed in the present document is numerically almost the same as the previous one, already listed in IRIS data base. Although, the upper bound on concentration of normal physiological "background" of methanol in humans has not been determined in this document, it seems that the exposure to methanol at the proposed RfD or RfC level may produce internal concentration not much different from the physiological background. This makes both, the existing and the proposed reference toxicity values for methanol very conservative. While the overall goal of developing reference toxicity values is to protect public health, perhaps, the revision of the current document would give an opportunity to U.S. EPA to derive RfC and RfD that would be not only health-protective, but also realistically close to the no-adverse-effect level in humans, with reasonable margin of safety and appropriate confidence.

COMMENTS SUBMITTED BY

David C. Dorman, Ph.D., DABT

Peer Review Meeting of EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Comments Submitted by Dr. David C. Dorman

(A) Toxicokinetics and PBPK Modeling

A1. Please comment on the scientific soundness of the PBPK model used in this assessment. The US EPA should be commended for developing the PBPK model and applying it to derivation of an RfC.

There are several weaknesses associated with the use of the model at this time including: (a) lack of external peer review through publication of the EPA PBPK model in the open scientific literature; (b) the model does not explicitly consider gestational/lactational exposure and compartments; (c) there is no clear rationale provided as to when the model should be applied (e.g., prior to the dosimetric extrapolations or apply the PBPK model to a POD determined through the BMD approach); and (d) the model does not include any a description of metabolism to formaldehyde or formate. These weaknesses do not prevent the EPA from using the model.

The model structure developed by the US EPA is based upon published models that were then adapted by the US EPA. The model structure is sound (although the use a bladder compartment is atypical – the EPA should consider recoding the model to include akidney/renal compartment that considers excretion of methanol by the kidney). The documentation provided by EPA is strong; however, the appendix describing the model contains extraneous information including email communications between scientists that do not contribute to a clear understanding of the model structure.

Page 3-45 (andelsewhere) includes a description of the two divergent models that were considered (Michaelis-Menten or not) which I found confusing. The EPA should more clearly describe their reasons for developing and using the models.

Page 3-49 – the description of the chamber volume should be expanded. I assume the equipment in question is the caging. Is that correct? Is there any evidence that incomplete mixing occurred since this group did examine methanol concentrations in different chamber locations prior to study start?

Page 3-50. The EPA has not clearly articulated why two different fractional absorption values were used based on the same data base (see pages 3-50 (60%) and 3-42 (86.5%)).

Page 3-50 – why was the second trimester group considered the most representative? This statement needs justification.

A2. Please comment on the scientific justification for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks.

The available pharmacokinetic datasets are quite variable – as correctly noted by the EPA in some cases investigators reported "corrected" blood methanol concentrations (where baseline concentrations were subtracted) while in other cases this information was provided. EPA states that the impact of including endogenous methanol was minimal – that is likely correct until one approaches low blood methanol concentrations seen with exposures approaching the proposed RfC. In this case these concentrations approach or exceed the contribution that results from the additional exogenous methanol exposure.

Ideally, the PBPK model should be revised to include endogenous methanol production/levels.

A3. The PBPK modeling effort assumed similar methanol pharmacokinetics between pregnant and non-pregnant animals. Please comment on the adequacy of the dose-metric extrapolation based on a PBPK model for non-pregnant adults (i.e., no fetal compartment) for predicting risks associated with fetal/neonatal brain concentrations of methanol.

The limited pharmacokinetic data suggests that this assumption is valid. However, there remains a question as to whether this model can be applied to neonatal rats where blood methanol concentrations are > 2-fold higher than those seen in dams under similar exposure conditions. This issue is important since the critical study used by EPA to derive an RfC involved combined gestational and lactational (inhalational) exposure of neonates. The use of an adult-based PBPK model could under predict potentially 'toxic' blood methanol concentrations. Indeed, the RfC estimate may more closely approximate that obtained using a more standard approach that doesn't rely on a PBPK model.

A4. EPA assumes limited methanol metabolism in the fetus because of limited alcohol dehydrogenase (ADH) activity in the human fetus, limited catalase and ADH activity in fetal rodents, and existing pharmacokinetic data that show nearly equal concentrations in maternal blood vs. the fetal compartment.Please comment on the validity of this assumption given the lack of data regarding potential alternate metabolic pathways in the fetus.

This assumption appears to be valid based on some of the existing methanol pharmacokinetic data (where maternal and fetal methanol concentrations are assessed). However, recent data published by Miller and Wells (2011) demonstrates that the embryotoxicity of methanol in cultured mouse fetuses is influenced by fetal catalase activity. In this study, methanol was more embryopathic in acatalasemic (aCat) mouse embryosthan their wildtype controls, with reduced anterior neuropore closure and head length only in catalase-deficient embryos. In concert with work published by Sweeting et al (2011) draw into question whether fetal methanol concentrations are a good predictor of teratogenic responses in different species – this data begs the question of which animal model(s) should be used in the methanol risk assessment where reproductive outcomes are of concern.

A5. Please comment on the scientific justification of the extrapolation approach from rats to humans for in-utero and neonatal lactational and inhalation exposures.

The use of the PBPK model is scientifically defensible and is a considerable strength of using a PBPK model. However, a concern remains as to whether the use of rodent data is appropriate since the metabolic activities and elimination kinetics of methanol in rodents are quite different from that seen in primates. As mentioned earlier, the model should be modified to include gestational and lactational components.

(B) Inhalation Reference Concentration (RfC) for Methanol

B1. A chronic RfC for methanol has been derived from a perinatal inhalation study of the effects from exposing rat dams and pups to methanol during gestation and lactation (NEDO, 1987).Reference values from mouse (Rogers et al., 1993) and monkey (Burbacher et al., 1999; 2004) developmental studies, were also derived and discussed, but were not chosen for the RfC. Please comment on whether the selection of the principal study has been scientifically justified.

Concerns remain concerning the use of the rat perinatal methanol study to derive the RfC. The endpoint of concern in this study was a decrease in neonatal brain weight (an effect also seen in a gestational only exposure albeit at a higher exposure dose). This response has not been replicated in other studies. Moreover, the analysis provided by the NEDO authors showed a gender difference (effects seen in males but not female rats). Moreover, the NEDO study relied on multiple t-tests as opposed to a more appropriate use of an ANOVA to evaluate gender and treatment responses. It is this reviewers understanding that this concern may not be relevant to EPA since EPA performed an independent BMD analysis of the data and demonstrated statistically significant trends. This should be more explicitly stated by EPA to alleviate concerns about the NEDO study.

The EPA did provide alternative RfC calculations which were appropriate.

B2. Reduction of brain weight at 6 weeks postnatally as reported in the NEDO (1987) developmental rat study was selected as the critical effect.Please comment on whether the rationale for the selection of this critical effect has been scientifically justified.Please identify and provide the rationale for any other endpoints (e.g., other reproductive and developmental effects reported in mouse and monkey studies) that should be considered in the selection of the critical effect.

There remains a general lack of transparency in the selection of this critical study and this single time point as the point of departure – this reviewer has the impression that the selection criteria used by EPA to determine the "best" study was the one that led to the lowest RfC. Although this is an appropriate approach (precautionary principle) the NEDO study remains problematic. Concerns about the statistical methods used by NEDO were raised by the external peer reviewers that reviewed this document for EPA; however, this concern does not appear to have been considered by the US EPA. I am also concerned that the EPA did not consider the full database from the NEDO study. Again, they arbitrarily considered only one time point (6 weeks) solely because it yielded the lowest value. This approach weakens the potential statistical power for a response that appears stable over a wide range of time points (3 to 8 weeks). In some ways, the NEDO study is the weakest of the three options. Although brain weight was evaluated there was a lack of

histological or functional follow-up for this response. The Burbacher study uses the most appropriate species (monkey) and examined a wide range of reproductive and neurotoxicologcal endpoints and significant pharmacokinetic data. The Rodgers study has undergone independent peer review, documents responses reported by other laboratories, and has quite robust group sizes.

B3. Benchmark dose modeling of decreased pup brain weight relative to maternal internal methanol doses predicted by the PBPK model was used to derive the point of departure (POD) for the RfC.Has the BMD/PBPK approach been appropriately conducted?Has adequate justification been provided for the selected internal dose metric, i.e., area under the curve (AUC) for methanol, in the blood of dams?Please identify and provide the rationale for any alternative approaches for the determination of the POD, including choice of another dose metric (e.g., methanol metabolized), and discuss whether such approaches are preferred to EPA's approach.

BMD modeling: "No response as this is outside my area of expertise."

The EPA could have considered either AUC or Cmax as the internal dosimetric of interest. The Agency has not adequately explained its rationale for the use of AUC rather than Cmax (e.g., see literature related to methanol and 2-methoxyethanol). I endorse the use of blood methanol as the dosimetric of interest.

I am confused by the rationale used by EPA to calculate the AUC. Table 5-2 indicated that the AUC was calculated with a 5 day 22 hr/day simulation. Why was a 5 day exposure duration was used?

B4. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the POD for the derivation of the RfC.It is assumed that these UFs account for variability in methanol dosimetry among human newborns following gestational and lactational exposure, and for uncertainty regarding the ratio of newborn-dose to maternal-dose in humans. Please comment on these assumptions and on the scientific justification for the selected UFs.

The UFs are poorly justified by the EPA. This is magnified when one reviews the alternative RfCs derived by EPA (Table 5-4)/ For example, it is surprising that the EPA used the same interspecies UF for rodent and nonhuman primate studies – given the fact that significant species difference exist between rodents and humans and less so between monkeys and people (use UF = 1). The database UF is poorly justified – the EPA continues to use this UF even in the face of a very rich toxicology database. When, if ever, will the database be adequate? The EPA is also basing the POD on the most sensitive studies that were conducted in neonatal animals from multiple species. If the critical study used is in neonates then why is an additional UF of 3 needed to account for children as a susceptible subpopulation. Again these UFs might be justified but they are poorly justified. Likewise, the discussion of possible gender effects (page 5-19) doesn't clearly articulate whether the experimental designs AND statistical analyses were adequate to determine whether or not a gender difference existed.

Just as importantly, the EPA systematically chose the most conservative approaches in developing the RfC. For example, they used a single SD for BMDL rather than a 4 or 10% changes as commonly used in some noncancer risk assessments (e.g., see page 5-23). These decisions were driven by the goal of obtaining a lower RfC value. There is a lack of transparency in this approach. In aggregate this yields an RfC value that

(C) Oral Reference Dose (RfD) for Methanol

C1. EPA concluded that the oral RfD should be derived using a route-to-route extrapolation from the more extensive inhalation database given the paucity of oral toxicity data.Please comment on whether the rationale for this approach has been scientifically justified and clearly explained. Please identify and provide the rationale for any alternative approaches for the determining the RfD and discuss whether such approaches are preferred to EPA's approach.

Theoretically this approach appears appropriate since there is limited data to consider that the route of exposure influences methanol disposition once this alcohol is absorbed. The US EPA should provide alternative RfC estimates that would be derived using traditional approaches.

C2. A PBPK model was used to derive the RfD via a route-to-route extrapolation, in which the internal-dose POD used for the derivation of the RfC based on data from the NEDO (1987) study was extrapolated to human oral exposure levels using the human PBPK model.Please comment on whether the rationale for this approach has been scientifically justified.Has adequate justification been provided for the selected internal dose metric, i.e., AUC for methanol, in the blood of dams?Is the PBPK model suitable for extrapolation of fetal and neonatal endpoints to human oral exposures?Please provide a detailed explanation.

As mentioned earlier, the model should include gestational and lactational components. This remains a weakness of the EPA approach.

C3. EPA applied the same UFs to the POD for the derivation of the RfD as for the RfC.Please comment on the rationale for the selection of the UFs.

The database for oral and inhalation are very different. Using an identical database UF is unexpected.

(D) General Charge Questions

D1. Is the Toxicological Review logical, clear and concise?Has EPA clearly synthesized the scientific evidence for non-cancer hazards?

The document has a number of significant weaknesses. Many of these were raised with the draft IRIS assessment for formaldehyde (NAS 2011). These include:

- Extensive reliance on a narrative approach with significant repetition of information throughout the document.
- There is a lack of inclusion and exclusion criteria. For example, several studies that could be supportive are lacking from the document – e.g., Bolon et al (1993, 1994) There are also other studies, including work in monkeys, with aspartame that may be supportive (e.g., Reynolds). Since Table 3-2 includes results from aspartame exposure this does not seem to be a clear exclusion criterion. Likewise, search terms and databases examined have been poorly defined.

- The document becomes extremely speculative. One of the more problematic section(s) related to the possibility that formaldehyde is the teratogen involved. Granted formaldehyde is more embryotoxic than either methanol or formate (on an equi-molar basis) this finding may or may not be related to the spectrum of teratogenic responses seen in animals (e.g., cervical rib anomalies, exencephaly, etc). This is especially important since there is a paucity of in vivo data suggesting that formaldehyde is teratogenic in animals (or people) (NAS, 2011). Another example relates to the discussion of parkinsonian signs in methanol-exposed people. These effects are commonly seen in hypoxic brain conditions so the question remains whether or not these represent a primary response or a secondary effect.
- The US EPA should rely more heavily on tables and not replicate repeatedly the same description of individual research studies.
- The Appendices are extremely difficult to read – there is an enormous amount of extraneous information provided.
- Table 3-3 should include the Dorman cynomolgus monkey study with a clear indication that it involved lung only exposure of anesthetized monkeys.
- In multiple locations the EPA mentions the results of the Fagan test performed by Burbacher and coworkers in monkeys exposed perinatally to inhaled methanol. At times, these results are used to support the selection of the NEDO rat study as the critical study. However, this effect lacked statistical significance in these animals. Is the EPA concluding that the effects seen were biologically significant despite the lack of a statistically significant response?
- I found section 3.4.2.4 confusing. There are other models that have been developed (Yoon et al., 2010, 2011) with inhaled manganese that could form the basis for a gestaional and lactational model.
- The discussion of a two compartment stomach (page 3-28 and elsewhere) for rodents need additional justification (squamous and epithelial portions?). Is this structure appropriate for people (as indicated on page 3-51).
- The EPA uses terms that describe model fits as “quite poor” (e/g/. see page 3-40 and elsewhere). This is at best a qualitative term that needs to be better clarified (visual inspection, goodness of fit, other?).
- I also found the use of different units of measure (e.g., mg/dL, mM, or ppm versus mg/m^3) frustrating. It would be ideal to pick one set of units (ppm would be preferred until calculation of the actual RfC value).
- Page 4-7 (and possibly earlier) the use of alcohol dehydrogenase inhibitors as a clinical ‘antidote’ should be discussed here. Many readers may not be familiar with this treatment approach.
- Page 4-18 – change to uriniferous tubule.
- Page 4-40 – does folate deficiency affect methanol concentrations significantly? Which data support this conclusion?

- Page 4-48 – change combustion product to metabolite.
- Table Legends should include exposure duration as well as exposure concentrations.

D2. Please identify any additional studies that would make a significant impact on the conclusions of the Toxicological Review and should be considered in the assessment of the non-cancer health effects of methanol.

- The Toxicological Review has incompletely considered the rabbit studies published by Sweeting and coworkers. These studies were discussed in an amended document provided to this reviewer well after the June 17 draft was provided. These studies were not considered in the EPA's consideration of inter-species differences (i.e., are rat or mice studies appropriate). AS noted earlier, there is a lack of inclusion and exclusion criteria. For example, several studies that could be supportive are lacking from the document – e.g., Bolon et al (1993, 1994) There are also other studies, including work in monkeys, with aspartame that may be supportive (e.g., Reynolds). Since Table 3-2 includes results from aspartame exposure this does not seem to be a clear exclusion criterion. Likewise, search terms and databases examined have been poorly defined.
- As noted by EPA the kinetics of methanol is heavily influenced by ventilation arte (page 3-18 and elsewhere). Leavens et al (2006) published data showing longitudinal changes in respiratory rate in pregnant rats that could be used with a gestational PBPK model.
- Dorman and coworkers also measured deciduas methanol concentrations. This could be considered in the discussion that occurs on page 3-52.
- The ethanol teratology literature has been largely ignored despite seom similarities in teratogenic response. This larger literature may help inform the MOA discussions in the draft document and help guide whether formaldehyde should be considered as the proximate teratogen.

D3. Please discuss research likely to substantially increase confidence in the database for *future* assessments of methanol.

Development of a PBPK model that considers gestation and lactational exposure. There is a need to replicate the findings of the critical study used by NEDO including the inclusion of additional neuropathological and neurobehavioral assessments. Although additional monkey studies could be considered the Burbacher study is extremely robust and should receive more attention by EPA. Additional MOA data – especially studies designed to resolve whether formaldehyde is involved in the developmental effects seen following peri-natal methanol exposure are warranted. These studies should include the use of dual labeled material to confirm fetal exposure.

Completion of surveys to examine blood methanol concentrations in the US population.

Improved understanding of the mode of action of methanol and it's metabolites (formaldehyde and formate) in the teratogenic response seen in animals.

COMMENTS SUBMITTED BY

Kenneth E. McMartin, Ph.D.

Peer Review Meeting of EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Comments Submitted by Dr. Kenneth E. McMartin

(A) Toxicokinetics and PBPK Modeling

A1. Please comment on the scientific soundness of the PBPK model used in this assessment.

The PBPK model that has been developed for this assessment is generally sound and has been thoroughly explained with generally appropriate justifications. An ideal scientific model of methanol pharmacokinetics/pharmacodynamics would have incorporated estimates of formate accumulation using parameters that estimated metabolism of formaldehyde (to formate and "other") and of formate (to CO_2 and urine formate excretion), similar to what was described for the Bouchard model. Incorporation of such parameters would be needed to help explain changes in blood methanol levels (since over long exposures, the rate of methanol elimination is governed by its metabolism). This, of course, would have made for a pretty complicated model. Because the EPA PBPK model is designed primarily to predict methanol blood levels across species (i.e., for HEC and HED calculations) for risk assessment purposes, the model presented in the document does a sufficient job.

An interesting aspect of the developed model is that it has employed various parameters in order to maximize the fit to existing methanol blood level data and some of these parameters don't make physiological or biochemical sense. At first glance, the use of a saturable term for stomach absorption of methanol seems peculiar since the absorption most likely is a passive diffusion, first order process. However, absorption of methanol could become saturated at very high concentrations simply because these high levels slow down gastric emptying (like ethanol is known to do), thus limiting absorption in the intestine. Hence, the need for the saturable term can be understood. It is also strange that a bladder component would be needed for humans and not for rodents. I realize that the authors were attempting to model urinary methanol excretion since such data are available for human studies (and not in the published rodent studies), but it would be interesting to see how the model were changed if the same bladder component were included in the rodent models.

The treatment of metabolism of methanol in the model is also hard to understand physiologically. For example, it is not clear why two saturable metabolic pathways are needed for the Sprague-Dawley rat and only one for the F344 rat in the sense that similar enzyme systems presumably operate in the two strains. It is understandable why the human model only incorporates one saturable pathway, because, at the low blood levels involved in this risk assessment, humans metabolize methanol exclusively by hepatic alcohol dehydrogenase (ADH) activity. Another odd result in the models is the apparent value of the biological constants (esp the K_m) that are calculated for the models (Table 3-10 and 3-11). These biological "constants" are actually calculated by the model through the curve fitting, that is the constants are changed iteratively in order to fit blood or urine level data from various published studies. The oddities lie in the fact that the resulting K_ms for the human and the rat range from 6 to 65 mg/L, whereas most biochemical studies with the rat and human ADHs report K_ms with methanol as a substrate in the range of 160 to 640 mg/L *in*

vitro (which are similar to the values reported for primates *in vivo* in Table 3-11). It is understandable that the "constants" are varied to fit the model, but the constants don't seem to reflect the true Michaelis values of the metabolic enzymes themselves. In both these situations, it appears that the mathematics of the model is driving the biology, rather than having the biological explanations first, with the math following behind.

A2. Please comment on the scientific justification for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks.

The model has used background subtraction of endogenous methanol from the animal studies to prepare the PBPK calculation of HECs. By and of itself, the subtraction of background levels in the application of the PBPK model for calculation of risk assessment numbers does not appear to substantially affect the numbers obtained. However, consideration of whether to subtract background levels is extremely important in how the results are applied – in fact the way that this risk assessment treats endogenous and exogenous methanol levels is highly questionable. The assessment assumes that the endogenous levels of methanol (and its metabolites) do not contribute to the formation of adverse effects, which presumably is true. Although this assumption is scientifically justified, it creates a major problem for risk assessment of substances like methanol that are found endogenously. Basically, as is shown by the resulting RfC and RfD that are determined in this document, exposures of humans from the levels of methanol at the RfC or the RfD produce no increase in blood methanol above the endogenous background. If endogenous levels of methanol do not contribute to adverse effects and an exposure does not produce an increase above background levels, how can that exposure lead to an adverse effect? The conundrum occurs because the PBPK model itself has built-in conservatism, the BMD calculation has built-in conservatism and then a 100-fold uncertainty is applied. All of these factors contribute to bring the "RfC/RfD exposure" down to the levels where there is essentially no exposure-induced increase in methanol levels above the endogenous, background level, which means there is essentially no risk.

A3. The PBPK modeling effort assumed similar methanol pharmacokinetics between pregnant and non-pregnant animals. Please comment on the adequacy of the dose-metric extrapolation based on a PBPK model for non-pregnant adults (i.e., no fetal compartment) for predicting risks associated with fetal/neonatal brain concentrations of methanol.

Existing literature strongly supports the assumption of similar pharmacokinetics between pregnant and nonpregnant animals (the Pollack and Brower as well as the Burbacher studies). Furthermore, studies have also indicated that fetal levels of the two major methanol metabolizing systems (ADH and catalase) are very low compared to adult levels, indicating that fetal tissues do not substantially impact the pharmacokinetics of methanol. As such, the use of a PBPK model based on non-pregnant adult data for predicting risks related to fetal concentrations of methanol can be substantiated for a dose-metric extrapolation. From a biological perspective, this makes sense – methanol is distributed evenly among tissues related to water content and so the levels in the fetal tissues should be similar if not identical as levels in the maternal tissues. Also, elimination of methanol controls its pharmacokinetics in most circumstances and the main driver for methanol elimination in the pregnant animal/human is the metabolic elimination by the maternal liver. Hence, the fetal levels of methanol are controlled mostly by the maternal liver, so the PBPK model is justified in utilizing nonpregnant data.

A4. EPA assumes limited methanol metabolism in the fetus because of limited alcohol dehydrogenase (ADH) activity in the human fetus, limited catalase and ADH activity in fetal rodents, and existing pharmacokinetic data that show nearly equal concentrations in maternal blood vs. the fetal compartment. Please comment on the validity of this assumption given the lack of data regarding potential alternate metabolic pathways in the fetus.

The assumption of limited methanol metabolism in the fetus is probably justified based on the existing studies showing low levels of ADH and catalase in fetal tissues. However, these studies have technically measured these proteins using indirect measures such as immunoblotting showing protein amounts or activity measures with ethanol as the substrate. Ideally an activity measurement using methanol as the substrate would be needed to confirm the low activity of methanol metabolism in fetal tissues. Nevertheless, the assumption that the fetal tissues do not substantially impact the pharmacokinetics of methanol is likely a good assumption.

A5. Please comment on the scientific justification of the extrapolation approach from rats to humans for in-utero and neonatal lactational and inhalation exposures.

Despite the various caveats noted above (A1-A4), the extrapolation conducted in this assessment from rats to humans is as good as can be done at present.

(B) Inhalation Reference Concentration (RfC) for Methanol

B1. A chronic RfC for methanol has been derived from a perinatal inhalation study of the effects from exposing rat dams and pups to methanol during gestation and lactation (NEDO, 1987). Reference values from mouse (Rogers et al., 1993) and monkey (Burbacher et al., 1999; 2004) developmental studies, were also derived and discussed, but were not chosen for the RfC. Please comment on whether the selection of the principal study has been scientifically justified.

Based on the analysis provided in the document and the results of the three studies themselves, the inhalation study of NEDO would appear to be the most justifiable of the three choices as the principal study. One advantage of the 1 generation or 2 generation NEDO rat studies over the others is the nearly continual exposure (20-22 h per day depending on the study) represents the types of exposures relative to the RfC/RfD (i.e. the daily exposure over the lifetime), whereas the mouse and monkey exposures were more like an occupational situation. A negative aspect of the NEDO study is that the critical effect (decreased brain weight) has not been reported in other studies, nor were there any corroborating clinical or pathological observations of depressed CNS activity noted in the rats in the NEDO study; in contrast, the critical effects in the Roger study (cervical ribs and CNS abnormalities) have been reported in other studies.

B2. Reduction of brain weight at 6 weeks postnatally as reported in the NEDO (1987) developmental rat study was selected as the critical effect. Please comment on whether the rationale for the selection of this critical effect has been scientifically justified. Please identify and provide the rationale for any other endpoints (e.g., other reproductive and developmental effects reported in mouse and monkey studies) that should be considered in the selection of the critical effect.

Reduction of brain weight in the NEDO rat study seems to be the critical effect – although the choice of 6 week data for the analysis is not scientifically justifiable (just because it produces the lowest BMD may be "standard procedure", but is not scientific). The biggest problem with this choice as the critical effect is that the statistical analysis of the brain weights used by NEDO is wrong and the risk assessment has not acknowledged this error. According to the 1987 report on this study, NEDO utilized multiple t-tests to compare the results between the various groups. Considering that there were four treatment groups (control and 3 levels of methanol exposure), the analysis should have involved an analysis of variance (ANOVA) to determine if there were an overall effect on brain weight, followed by an appropriate post-hoc test to examine for differences among groups (in order to test if one of the levels of methanol was significantly different from control). This document shows the NEDO data in Table 4-10 with the footnote that the statistical differences were calculated by the authors. A re-analysis using ANOVA would seem to be appropriate to ensure that the brain weights are significantly reduced by a methanol treatment, thus allowing for the subsequent BMD analysis. The 10-15% decrease in brain weight in males is probably "significant", so this might be a moot point, but a re-analysis is necessary to determine which time frame and which methanol level are used for the BMD analysis. This issue is of greater importance because the NEDO study did not report (examine for?) any corroborating information including effects on the brain such as clinical signs/symptoms or CNS pathology, as noted above.

B3. Benchmark dose modeling of decreased pup brain weight relative to maternal internal methanol doses predicted by the PBPK model was used to derive the point of departure (POD) for the RfC. Has the BMD/PBPK approach been appropriately conducted? Has adequate justification been provided for the selected internal dose metric, i.e., area under the curve (AUC) for methanol, in the blood of dams? Please identify and provide the rationale for any alternative approaches for the determination of the POD, including choice of another dose metric (e.g., methanol metabolized), and discuss whether such approaches are preferred to EPA's approach.

The BMD/PBPK approach has been conducted appropriately for the most part (with the exception noted in B2). The selected internal dose metric (methanol AUC) is preferable over the alternative of methanol metabolized. Although there are studies showing that formaldehyde (HCHO) is the most "toxic" of the three compounds (methanol, formate and HCHO) in whole embryo cultures, this is a common situation with HCHO since it is a very interactive aldehyde. The key question is whether any of the metabolites are transported into the fetus or whether they are formed in the fetus (and if so, reach the fetal brain). Most of the existing studies have ruled out a role for formate in these developmental effects, so the only question is the potential role of formaldehyde. Certainly based on all existing studies, HCHO is not formed in the dam/mother and transported into the fetus. Also, fetal metabolism of methanol to HCHO appears to be minimal if at all. Even if a small amount of HCHO were generated in the fetal liver, it would not be transported to the brain (because it would be rapidly metabolized to formate within the fetal liver cell or would rapidly bind to components in the cell – either way it is not likely to even leave the liver cell).

The increased effect seen when methanol is administered to glutathione (GSH)-depleted animals does not necessarily imply that the MOA for methanol involves metabolism to HCHO (which has been suggested because GSH depletion should decrease HCHO elimination allowing for higher HCHO levels). Depletion of GSH, as the major cellular antioxidant, will also increase the accumulation of reactive oxygen species (ROS) – since generation of ROS is a viable mechanism by which methanol induces its effects, the increase in ROS may explain the increased effects of methanol in GSH-depleted animals.

The existing studies do indicate that methanol itself is the responsible agent (although possibly through generation of ROS by unknown mechanisms). Hence, the dose metric is either methanol AUC or Cmax. The justification for using AUC is tenuous – there are some studies suggesting that duration of exposure is important, but the fact that the effect on brain weight does not differ between the 3, 6 and 8 week periods does conflict with this assumption. Use of Cmax would suggest that there is a threshold response, i.e. a certain blood level must be reached. This alternative is attractive (formate Cmax is the key to ocular toxicity for example).

B4. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the POD for the derivation of the RfC. It is assumed that these UFs account for variability in methanol dosimetry among human newborns following gestational and lactational exposure, and for uncertainty regarding the ratio of newborn-dose to maternal-dose in humans. Please comment on these assumptions and on the scientific justification for the selected UFs.

The 10-fold uncertainty factor for inter-human variability is generally justified by likely genetic variations in methanol metabolizing enzymes (thus, producing differences in methanol elimination) and by highly likely differences in human nutrition (folate deficiency is known to exacerbate the developmental effects). The 3-fold uncertainty factor for the pharmacodynamic human-to-animal extrapolation is also likely justified based on the lack of existing knowledge regarding the mechanism of the developmental effects – hence not enough information to assess pharmacodynamic differences in animals vs. humans.

The database uncertainty factor of 3 is not at all justified. Based on standard procedures, there is never enough data to be certain regarding a risk "assessment" (that is why it is called risk assessment not a risk determination). More importantly, conservative assumptions are always used on all these procedures (PBPK assumes the most conservative scenarios, BMD analysis itself favors the conservative numbers and lastly when given the choice of alternative BMD numbers such as those obtained from the 3 vs. 6 vs. 8 week data, the lowest, i.e. most conservative, number is chosen. Thus, because of the conservative approach in this risk assessment and because of the fact that methanol has an endogenous profile, there is no reason for the additional UFd of 3.

(C) Oral Reference Dose (RfD) for Methanol

C1. EPA concluded that the oral RfD should be derived using a route-to-route extrapolation from the more extensive inhalation database given the paucity of oral toxicity data. Please comment on whether the rationale for this approach has been scientifically justified and clearly explained. Please identify and provide the rationale for any alternative approaches for the determining the RfD and discuss whether such approaches are preferred to EPA's approach.

The existing oral studies do not appear to be adequate for determination of an RfD. Either no dose-response effects were reported in the studies (Soffritti) or the data were not suitable for BMD determination (EPA 86). Because of the thoroughness of the PBPK model as it is currently developed, the route-to-route extrapolation from inhalation data to estimate the oral RfD is a logical approach. The explanation and justification is adequate.

C2. A PBPK model was used to derive the RfD via a route-to-route extrapolation, in which the internal-dose POD used for the derivation of the RfC based on data from the NEDO (1987) study was extrapolated to human oral exposure levels using the human PBPK model. Please comment on whether the rationale for this approach has been scientifically justified. Has adequate justification been provided for the selected internal dose metric,i.e., AUC for methanol, in the blood of dams? Is the PBPK model suitable for extrapolation of fetal and neonatal endpoints to human oral exposures? Please provide a detailed explanation.

Despite the comments on the PBPK model in Section A above, the developed model is useful for the extrapolation from the NEDO inhalation study to produce an oral POD. As in C1, the rationale and justification is adequate. Despite the various caveats noted in Section B, methanol AUC is suitable.

C3. EPA applied the same UFs to the POD for the derivation of the RfD as for the RfC. Please comment on the rationale for the selection of the UFs.

The same comments regarding UFs in section B4 can be applied here. Just to reiterate, the UFd of 3 is not justifiable and should be eliminated.

(D) General Charge Questions

D1. Is the Toxicological Review logical, clear and concise? Has EPA clearly synthesized the scientific evidence for non-cancer hazards?

In general the Toxicological Review is thorough in its description of the numerous studies as well as of the BMD and PBPK modeling. Two general comments – 1) it is somewhat repetitious, which makes it difficult to appreciate the key points (having to wade through the extraneous material also); 2) the inclusion of the discussion of the CNS effects produced by acute methanol overdosing is not appropriate – it seems in a biased way to validate the subsequent choice of the NEDO study (decrease in brain weights suggesting a methanol-induced CNS effect). In reality, the two effects are not related (the acute methanol CNS pathology occurs in the exposed subject per se, due most likely to formate accumulation, while the reduced brain weight is a developmental effect unrelated to formate).

The following minor comments on the document are noted.

1. Table 3-5 – why are the ranges of blood methanol for the Stanton study expressed downwards (from high to low), while the other ranges go from low to high? IS it true that the blood formate values reported by Horton are identical for the F-344 rats as for the monkeys (Table 3-3)? The last entry in Table 3-5 for Lee et al is not clear – the numbers for methanol and formate do not align with the various dose groups so it is hard to tell which is which.
2. Section 3.4.2.1 is identical to section 3.4.2.3 – one can be deleted.
3. Table 3-11 and p. 3-40. The data are cited as mean $\pm$ SE in text but as SD in the table.
4. p. 4-2. The reports by Bennett et al. and by Benton/Calhoun are really from the same study of the same epidemic. This should be clarified here (or one reference eliminated).
5. p. 4-6. As written, it sounds like there have only been two cases of inhalation/skin exposure. Not true, there have been many published reports, although mostly by Woods and coworkers in the early 1900s.
6. p. 4-14. I realize the data quoted come from an encyclopedic listing of LD50s, but as listed, the data imply that the methanol is not very lethal to the monkey. IN fact, monkeys are quite susceptible to the acute oral toxicity of methanol, with lethality occurring in the 3-4 g/kg range which is much less than in the rodent. A caveat should be added to indicate the unique sensitivity of the primate to methanol.
7. p. 4-70 – it should be Table 9 not 10 in the first line. Also in line 19 the referencing is misplaced. Johlin et al studied hepatic folate levels between species and not anything to do with catalase levels (so reference should be in the previous sentence)
8. p. 83 line 19, Latter not former.

D2. Please identify any additional studies that would make a significant impact on the conclusions of the Toxicological Review and should be considered in the assessment of the non-cancer health effects of methanol.

In the Addendum to the Draft Peer Review, there is discussion of interesting new results by Wells and coworkers (2010 and 2011). These investigations have compared the metabolism and pharmacokinetics of methanol in mice, rabbits and monkeys. As would be expected from numerous previous studies, monkeys showed a markedly higher level of formate accumulation and a slower rate of methanol elimination compared to mice. The interesting observation is that the rabbit appeared to be somewhere in the middle. As seen in Figure 4 of the Sweeting 2010 paper, methanol is cleared more slowly in the rabbit such that blood levels remain elevated for nearly 40 h (compared to only 20 h in mice). Similarly, formate blood levels are elevated for 40 h in rabbits and only 20 h in mice. The authors interpreted these data to indicate that methanol kinetics in rabbits more closely approximates those in primates (hence humans), so rabbits might be a better animal model than mice (rodents) for developmental studies. However, it is important to note that the peak formate level in the rabbit is no higher than that in mice, just the duration – the duration of formate elevation is strictly controlled by the elimination of methanol (once the latter is gone, formate is gone also). In that sense the rabbit is still different from the primate in terms of formate accumulation. However, formate doesn't appear to play a role in the developmental effects of methanol. As such, the

slower elimination of methanol in the rabbit would make it a preferred model over mice for developmental studies – however, since the key study used for this RfC/RfD analysis is the NEDO rat study, the key comparison would be between rabbits and Sprague-Dawley rats (which also happen to eliminate methanol more slowly than do mice).

The Discussion of these studies as well as the other ones by Wells and coworkers (esp the publication regarding the role of ROS in mediating the effects of methanol) needs to be improved in the document. Although the "results" of these studies are presented in the Addendum, the ramifications of these studies is not presented well or at all in the sections related to choice of POD, critical effect, etx.

D3. Please discuss research likely to substantially increase confidence in the database for future assessments of methanol.

Ideally, a study that fully characterizes methanol metabolism in the intact fetus and the dam using the rat as model would allow for a direct assessment of the role of fetal ADH and/or catalase in producing HCHO in the fetus (as opposed to the existing studies that only assess protein levels or activities using ethanol as substrate). Such a study would impact on the MOA for the developmental effects of methanol.

Studies of the role of ADH and catalase in the metabolism of methanol by F-344 and Sprague-Dawley rats would clarify why there might be two saturable pathways in one strain but only one in the other (as implied by the PBPK model). These studies should also be done in a way to estimate the Km and Vmax for methanol metabolism to further refine the PBPK model with defined rather than fitted parameters.

There is some concern that the NEDO rat studies are the only one to have reported decrease brain weight as a developmental effect of methanol. There are other studies showing various other effects of similar exposures to methanol, but none report decreased brain weight, nor any really dramatic CNS pathology (such as leading to real behavioral changes – the changes in several studies have been incredibly minor). Thus, it would be ideal to have an independent study to essentially repeat the NEDO-type exposures (i.e. 20+ hours per day instead of the 2.5 – 7 hours) to test if there are effects of methanol on the CNS under those circumstances (both on brain weight and on brain function/pathology).

Again ideally it would be good to have a well conducted oral developmental study of methanol in order to produce sufficient data for an RfD (instead of having to do extrapolations from an inhalation study).

Extra. Please comment on the proposed RfD and RfC values for their intended use in risk assessment. Are these numbers more conservative than they need to be to protect public health?

The RfD and RfC values have been appropriately derived based on the BMD/PBPK analysis utilizing "standard EPA procedures", but the resulting values lack scientific credence and are not logical in the sense of the exposures expected for humans. The problem scientifically is that the resulting values reflect a methanol exposure that does not increase the blood methanol concentration in the exposed human. Because of the background level of methanol in all humans lies in the range of 2 mg/L, the projected increase in

methanol level from the RfC/RfD exposure is only 0.04 mg/L, i.e. a level that is really indistinguishable from the background. The implications of this include that all humans would be susceptible to developmental effects of methanol no matter what exposure they had experienced – not a suitable endpoint for risk assessment. As stated above, presuming that endogenous levels of methanol do not contribute to adverse effects and an exposure does not produce an increase above background levels, how can that exposure lead to an adverse effect? The conundrum occurs because the PBPK model itself has built-in conservatism, the BMD calculation has built-in conservatism and then a 100-fold uncertainty is applied. All of these factors contribute to bring the "RfC/RfD exposure" down to the levels where there is essentially no exposure-induced increase in methanol levels above the endogenous, background level, which means there is essentially no risk. So in this case of an endogenous chemical, the numbers are more conservative than necessary.

COMMENTS SUBMITTED BY

Stephen M. Roberts, Ph.D.

Peer Review Meeting of EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Comments Submitted by Dr. Stephen M. Roberts

(A) Toxicokinetics and PBPK Modeling

A1. Please comment on the scientific soundness of the PBPK model used in this assessment.

Construction of PBPK models is outside my expertise. I see no obvious flaws in the model, but cannot comment on a technical level regarding its scientific soundness.

A2. Please comment on the scientific justification for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks.

The case for subtraction is based upon a stated goal of determining noncancer risk associated with exposures that increase body burden of methanol or its metabolites above endogenous levels. Two assumptions are stated: "(1) endogenous levels do not contribute significantly to the adverse effects of methanol or its metabolites; and (2) the exclusion of endogenous levels does not significantly alter PBPK model predictions." (pg 3-27). With respect to the second assumption, modeling with and without incorporation of background levels was tested using rat pharmacokinetic data with the stated result that incorporation of background had minimal effect (<1%) on the point of departure (POD). Given the doses of methanol used in the rat studies, the contribution of endogenous methanol to total blood levels was no doubt very small, and it is not surprising that the POD changed little whether or not endogenous methanol was included. The more important matter is the first assumption. If the proposed RfC and RfD values for methanol were associated with blood methanol levels much higher than endogenous levels, the contribution of endogenous levels would not be an issue. That is not the case, however. The RfC and RfD correspond to blood methanol concentrations in humans squarely in the range of normal "background" levels. I'm not aware of any evidence that endogenous and exogenous methanol are distinct in their potential to produce noncancer effects, and in fact there is no attempt to present a scientific argument on that point in the Toxicological Review. So, under the circumstances, the first assumption is not met.

A3. The PBPK modeling effort assumed similar methanol pharmacokinetics between pregnant and non-pregnant animals. Please comment on the adequacy of the dose-metric extrapolation based on a PBPK model for non-pregnant adults (i.e., no fetal compartment) for predicting risks associated with fetal/neonatal brain concentrations of methanol.

Studies presented in the Review indicate that the pharmacokinetics of methanol are similar in pregnant and non-pregnant rodents, providing a scientific basis for modeling methanol concentrations in pregnant animals based upon data from non-pregnant adults. Given the data currently available, I understand the rationale for omitting a fetal compartment in the PBPK model. However, I think that for PBPK modeling to be effective, a fetal compartment will ultimately be needed. Studies such as Sweeting et al. (2011) suggest that maternal blood methanol concentrations alone are insufficient to explain developmental toxicity from methanol, even

within the same species. PBPK modeling is most useful when the proximate form of the toxicant and mode of action are known, which is unfortunately not the case with developmental effects of methanol.

A4. EPA assumes limited methanol metabolism in the fetus because of limited alcohol dehydrogenase (ADH) activity in the human fetus, limited catalase and ADH activity in fetal rodents, and existing pharmacokinetic data that show nearly equal concentrations in maternal blood vs. the fetal compartment. Please comment on the validity of this assumption given the lack of data regarding potential alternate metabolic pathways in the fetus.

The assumption of limited methanol metabolism in the fetus is valid from the perspective of PBPK modeling of methanol concentrations. There is sufficient information to show that ADH and catalase metabolism of methanol are relatively low in both the rodent and human fetus. Significant alternative pathways of metabolism of methanol in the fetus have not been identified. That is not to say that fetal metabolism is insignificant from the standpoint of methanol developmental toxicity, however. Pathways in the fetus that are quantitatively minor compared with maternal metabolism can nonetheless be very important in determining adverse effects. For example, studies by Wells and Miller (2011) suggest that fetal catalase activity is important in determining susceptibility to methanol developmental effects in rodents.

A5. Please comment on the scientific justification of the extrapolation approach from rats to humans for in-utero and neonatal lactational and inhalation exposures.

As explained in Appendix C, experimental data indicate that the kinetics of inhaled methanol are similar in pregnant and non-pregnant mice, and that maternal blood and fetal methanol concentrations are approximately equal. This is assumed to apply to rats as well, and that maternal blood methanol concentrations are therefore appropriate indicators of fetal concentrations during gestation in this species as well. Because offspring of maternally exposed rats have subsequent exposure through lactation and direct inhalation, methanol concentrations in pups were likely greater than in the dams. Assuming that the same difference occurs in human mothers and offspring, this difference was deemed relatively inconsequential for the purposes of the analysis. The first assumption – similarity in maternal-fetal concentrations in both mice and rats – seems reasonable given similarities in the nature of distribution of methanol in the body across species. The second assumption requires a much greater leap of faith – that lactational and inhalation exposure postpartum in rats and humans are sufficiently similar that the same maternal/offspring methanol concentration ratios will be seen. The Review points out a study by Stern et al. (1996) indicating that methanol concentrations in pups exposed by inhalation were approximately 2.25 times higher than the dams, and states that a similar ratio probably occurred in the NEDO (1987) study used to generate the RfC, given similar designs of the studies. It is then assumed that the maternal/infant methanol concentration ratio in human infants would not be significantly greater than that observed in rats. This is purely speculation. Differences in exposure as well as methanol clearance between human infants and pups could lead to substantially different maternal/offspring methanol concentration ratios (higher or lower). This is a significant source of uncertainty in the extrapolation from rats to humans.

(B) Inhalation Reference Concentration (RfC) for Methanol

B1. A chronic RfC for methanol has been derived from a perinatal inhalation study of the effects from exposing rat dams and pups to methanol during gestation and lactation (NEDO, 1987). Reference values from mouse (Rogers et al., 1993) and monkey (Burbacher et al., 1999; 2004) developmental studies, were also derived and discussed, but were not chosen for the RfC. Please comment on whether the selection of the principal study has been scientifically justified.

Among the three developmental studies considered, only the Burbacher et al. studies in monkeys were discussed in terms of limitations. [Note: A number of the limitations in the Burbacher et al. studies are overstated, such as the inclusion of wild caught monkeys and the influence of C-sections on results.] The explanation for choosing the RfC from the NEDO study over the RfC from the Rogers et al. study is simply that the value is lower. Choosing the more lower, more conservative value is a policy choice rather than one based upon the scientific strengths of the two studies.

B2. Reduction of brain weight at 6 weeks postnatally as reported in the NEDO (1987) developmental rat study was selected as the critical effect. Please comment on whether the rationale for the selection of this critical effect has been scientifically justified. Please identify and provide the rationale for any other endpoints (e.g., other reproductive and developmental effects reported in mouse and monkey studies) that should be considered in the selection of the critical effect.

A significant reduction in brain weight on an absolute basis was reported in the NEDO (1987) developmental rat study. No abnormal brain histopathology or functional deficits were noted. A recent review of this study conducted for EPA raised questions about the statistical analysis and whether or not the small brain weight changes in fact are significant and represent an adverse effect (External Letter Peer Review of Reports Documenting Methanol Studies in Monkeys, Rats, and Mice Performed by the New Energy Development Organization (NEDO), Peer Reviewer Comments, June 16, 2009). In contrast, the Rogers et al. study is considered rigorous and well described (see, for example, NTP-CERHR, 2003), and the increases in cervical ribs and supernumerary ribs observed in this study could be considered a more scientifically justified critical effect.

B3. Benchmark dose modeling of decreased pup brain weight relative to maternal internal methanol doses predicted by the PBPK model was used to derive the point of departure (POD) for the RfC. Has the BMD/PBPK approach been appropriately conducted? Has adequate justification been provided for the selected internal dose metric, i.e., area under the curve (AUC) for methanol, in the blood of dams? Please identify and provide the rationale for any alternative approaches for the determination of the POD, including choice of another dose metric (e.g., methanol metabolized), and discuss whether such approaches are preferred to EPA's approach.

An explanation is provided in Appendix C why maternal AUC is considered to be an appropriate dose metric for offspring. There are two issues: 1) use of a maternal dose metric to represent offspring; and 2) choice of AUC versus another metric for internal dose. The first issue is discussed in the response to Question A5. With respect to the second issue, internal dose metrics can be selected based upon mechanistic or empirical considerations. Given that the mode of action of methanol developmental effects is unknown, empirical evaluation is left. Unfortunately, there is little in the way of data for empirical evaluation. Because there is evidence that effects on brain weight undergo some recovery when exposure is terminated,

this was viewed as indicating that both the level and duration of effect are important (see pg C-2). Because AUC incorporates a time component, and because it is commonly used as an internal dose metric, it was selected for this endpoint. AUC is a reasonable choice, providing a measure of the average concentration over time. Other metrics could be considered, but as noted above, there are no data with which to argue that any would be a better choice.

B4. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the POD for the derivation of the RfC. It is assumed that these UFs account for variability in methanol dosimetry among human newborns following gestational and lactational exposure, and for uncertainty regarding the ratio of newborn-dose to maternal-dose in humans. Please comment on these assumptions and on the scientific justification for the selected UFs.

A composite UF of 100 was selected based upon a UF_H of 10, a UF_A of 3, and a Database UF of 3. A UF_A of 3 was selected because PBPK modeled was thought to address the pharmacokinetic component of this UF. A Database UF of 3 was chosen based upon perceived deficiencies in the toxicity database, mentioning uncertainty regarding the most appropriate test species and limitations in developmental toxicity studies currently available. Selection of individual UFs appears to be consistent with contemporary EPA guidance and practice, although a strong argument could be made for eliminating the database uncertainty factor. As noted in the question, it is implied that the UFs cover specific uncertainties in derivation of the RfC, including variability in methanol dosimetry among human newborns following gestational and lactational exposure and the ratio of newborn dose to maternal dose in humans. Because of limited information available, it is difficult to judge the potential error associated with assumptions regarding human newborn exposure and resulting methanol concentrations. However, it seems likely that the error is not so great that a composite UF of 100 would be inadequate.

(C) Oral Reference Dose (RfD) for Methanol

C1. EPA concluded that the oral RfD should be derived using a route-to-route extrapolation from the more extensive inhalation database given the paucity of oral toxicity data. Please comment on whether the rationale for this approach has been scientifically justified and clearly explained. Please identify and provide the rationale for any alternative approaches for the determining the RfD and discuss whether such approaches are preferred to EPA's approach.

Limitations in existing oral studies are adequately described in Section 5.2.1, supporting a decision not to use oral study data to develop an oral RfD. Observations of effects in subchronic and chronic oral toxicity studies by U.S. EPA (1986) and Soffritti et al. (2002), respectively, were insufficient to support quantitative analysis to establish a NOAEL or BMD. Also, I agree with the decision not to use any of the oral methanol reproductive and developmental toxicity studies that used very high methanol doses over selected gestational periods. Parenthetically, I am surprised that better oral chronic toxicity data are not available for a chemical of the importance of methanol.

Under the circumstances, the decision to use route-to-route extrapolation and inhalation toxicity data is justified, and the rationale is clearly explained in the Toxicological Review (pg 5-23). I have no suggestions for an alternative approach to developing an oral RfD that would be preferable.

C2. A PBPK model was used to derive the RfD via a route-to-route extrapolation, in which the internal-dose POD used for the derivation of the RfC based on data from the NEDO (1987) study was extrapolated to human oral exposure levels using the human PBPK model. Please comment on whether the rationale for this approach has been scientifically justified. Has adequate justification been provided for the selected internal dose metric, i.e., AUC for methanol, in the blood of dams? Is the PBPK model suitable for extrapolation of fetal and neonatal endpoints to human oral exposures? Please provide a detailed explanation.

The route-to-route extrapolation is based on assumptions that the critical effect is related to methanol concentrations in blood and is independent of the route of exposure by which methanol reaches the blood. Both assumptions are reasonable. It follows that the POD and the dose metric should be the same whether from oral or inhalation exposure. The PBPK model is suitable to extrapolate an external human oral dose that corresponds to the internal dose POD. The appropriateness of the POD with respect to fetal and neonatal endpoints from oral exposure is the same as with inhalation exposure, i.e., has the same limitations (see answers to previous questions).

C3. EPA applied the same UFs to the POD for the derivation of the RfD as for the RfC. Please comment on the rationale for the selection of the UFs.

The Review states that because the same dataset, endpoint, and PBPK model were used to derive the RfC and RfD, the same UFs were applied. I agree with this rationale. As noted in the response to B4, it is questionable whether a database uncertainty factor of 3 is needed given the data available on sensitive endpoints in multiple species.

(D) General Charge Questions

D1. Is the Toxicological Review logical, clear and concise? Has EPA clearly synthesized the scientific evidence for non-cancer hazards?

The Toxicological Review is generally logical and clear, but not concise. The format of the document contributes to redundancies, and presentation of some topics is fragmented, forcing the reader to synthesize information presented in more than one section of the main document and appendices. This is problem inherent in the current format for IRIS toxicological reviews. Sometimes subtle points are lost in the repetition. A great deal of information in order for the analysis to be transparent, but this shouldn't get in the way of clearly highlighted key points and decisions. A different format could be much more effective in conveying critical information, interpretations, and decisions regarding available, relevant toxicological literature.

D2. Please identify any additional studies that would make a significant impact on the conclusions of the Toxicological Review and should be considered in the assessment of the non-cancer health effects of methanol.

A description of additional studies was added as an Addendum to the Toxicological Review provided to the Panel. To these studies, the paper by Miller and Wells (Toxicol. Appl. Pharmacol. 252:55-61, 2011) should be added. This paper is germane to the discussion of potential mode of action for methanol developmental effects and also extrapolation of observations in mice to other species.

D3. Please discuss research likely to substantially increase confidence in the database for *future* assessments of methanol.

Existing research clearly indicates the potential for methanol to produce developmental effects, but recent studies have suggested substantial differences among species and even strains, both in terms of susceptibility and type of effect. This creates enormous uncertainty in the extrapolation of effects from one species/strain to humans. Research to explain the basis for differences in species/strain developmental effects is essential. Also, sound dosimetry is compromised by lack of understanding of the proximate toxicant and mode of action for developmental toxicity. This must be resolved in order to more effectively utilize PBPK modeling for extrapolation across species.

E. Please comment on the proposed RfD and RfC values for their intended use in risk assessment. Are these numbers more conservative than they need to be to protect public health?

The process of developing these RfC and RfD values has produced a result that is counter-intuitive, implying that individuals with no unusual methanol exposure may be at risk of developmental effects. That's implausible, and clearly signals the need to re-evaluate how to consider background methanol concentrations in the development of credible toxicity values. What is missing from this and other IRIS toxicological reviews is an assessment, after going through the process of RfC and RfD development, whether the resulting values make sense, i.e., are they logical in the context of exposures and effects anticipated in human populations? For many chemicals, this may be impossible to judge; there simply is not enough information about exposure or exposed populations to make a determination. For more familiar chemicals, if there is some question about the validity of the toxicity values, it should incumbent upon the toxicological review to make a case why the toxicity values are reasonable. Public comments have argued that the proposed methanol RfC and RfD values correspond to doses commonly experienced by the public (e.g., one small glass of orange juice daily), and are projected to produce increases in blood methanol concentrations for most individuals that would keep them with the background range. The NTP-CERHR methanol panel considered common exposures to methanol and concluded that they pose no immediate concern for developmental toxicity (NTP-CERHR, 2003). This appears to be a reasonable conclusion and creates a real credibility problem for the proposed methanol RfC and RfD.

COMMENTS SUBMITTED BY

Andrew G. Salmon, Ph.D.

Peer Review Meeting of EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Comments Submitted by Dr. Andrew G. Salmon

(A) Toxicokinetics and PBPK Modeling

A1. Please comment on the scientific soundness of the PBPK model used in this assessment.

The model is based on extensive and well-documented knowledge of the metabolism and kinetics of methanol and related compounds, including investigations of the differences between experimental animal species and humans. It also builds on the experience gained from several previous investigations of methanol toxicokinetics, and successfully avoids some limitations and peculiarities of these earlier models in developing a generally applicable model framework. The Agency is to be particularly commended for developing a consistent model framework and sets of species-specific parameters which have been validated across several somewhat diverse data sets. This considerably increases the confidence which may be placed in its conclusions, and is a welcome improvement on some earlier PBPK modeling exercises which did not test their validity by attempting to fit multiple independent datasets. The ability to effectively model the limited, but relevant, toxicokinetic data in humans as well as in the two experimental animal species of interest increases confidence in the usefulness of the model for interspecies extrapolation.

One issue which needs attention is the fitting of the PBPK model to kinetic data for Sprague-Dawley (SD) rats. This is of significance for analysis of the NEDO developmental study. The analysis presented in the draft report relies on the dataset which appears most complete (which is an appropriate choice), but this results in an unexpectedly low value for the uptake fraction by inhalation for this strain of rat. It appears from discussions at the meeting that additional data from NEDO are now available which not only provide an alternative basis for parameterizing the model for SD rats, but also result in an uptake fraction similar to that seen for Fischer rats and mice, which is intrinsically more plausible biologically. This needs to be checked out in detail, and its consequences for the predicted HECs in the NEDO (1987) rat study explored.

A2. Please comment on the scientific justification for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks.

Conceptually this is a reasonable thing to do, since the non-cancer health effect data similarly consider the difference in incidence between controls (with the background methanol level) and exposed (with background plus exposure-related levels). Where there might be a problem with this is if the background levels were comparable to or larger than the exposure-related level, or if there was evidence of substantial perturbation of the background methanol metabolism by the additional exposure-related component. These concerns do not appear to be applicable at the exposure levels of interest in analysis of the experimental animal data and human studies analyzed for the development of the RfC. (The question of whether the background concentrations are comparable to levels achieved at the RfC/RfD is a different, and not necessarily relevant, question.)

In particular, alternative model analyses in which the background concentrations were explicitly included produced very similar results to those looking at additional levels only. Since both approaches provide reasonable (and essentially identical) fits to the actual incidence data it appears that there are no major interactions between background and exposure-related levels to complicate the analysis of the animal datasets. There are modest differences in HEC and HED predictions with background included or excluded. It is not apparent from the narrative either in the main text or the appendix describing the model why these differences arise. I did not find any analysis of whether these differences should be considered significant from a statistical standpoint: it may be that there are insufficient human data to evaluate this question. On balance the Agency's decision to exclude the background levels from the calculations seems reasonable, although some further explanation is desirable. Apparently models both with and without inclusion of background levels have been developed, so that specific differences in these approaches could be identified. Including the background levels in the models necessarily increases the model complexity and like any model enhancement may increase the uncertainty in the final result, especially when as in this case it may be difficult to design a test of its validity. Part of the problem is the considerable variability and uncertainty in human background values for methanol. Also there is significant uncertainty as to mechanism and the extent to which individual genetic variability or adaptation may affect the possible toxicological significance of those values.

A3. The PBPK modeling effort assumed similar methanol pharmacokinetics between pregnant and non-pregnant animals. Please comment on the adequacy of the dose-metric extrapolation based on a PBPK model for non-pregnant adults (i.e., no fetal compartment) for predicting risks associated with fetal/neonatal brain concentrations of methanol.

The report describes various practical investigations and studies with PBPK models which indicate that, due to the free and rapid distribution of methanol throughout most maternal and fetal tissues, inclusion of a fetal compartment in the model adds little to the prediction of fetal methanol levels. The decision to use the non-pregnant model is supported by the data, and well justified in the narrative.

A4. EPA assumes limited methanol metabolism in the fetus because of limited alcohol dehydrogenase (ADH) activity in the human fetus, limited catalase and ADH activity in fetal rodents, and existing pharmacokinetic data that show nearly equal concentrations in maternal blood vs. the fetal compartment. Please comment on the validity of this assumption given the lack of data regarding potential alternate metabolic pathways in the fetus.

It is obviously difficult to get objective measures of fetal metabolic capabilities, and although there are some data in this case there is inevitably a substantial measure of uncertainty involved. Some data described in the report (including the recent studies by Sweeting et al., 2010; 2011included in the addendum) suggest that although ADH1 levels in the fetus are typically low, this shortfall may be compensated by relatively higher catalase activity even in species where the latter route is not an important factor in the adult. However, the limited data available on methanol levels in the fetus relative to the maternal levels imply that fetal metabolic clearance is not sufficient relative to the rate of equilibration to substantially differentiate these two compartments, supporting the PBPK analysis described in the report.

A5. Please comment on the scientific justification of the extrapolation approach from rats to humans for in-utero and neonatal lactational and inhalation exposures.

Appropriate species-specific PBPK models have been developed, and validated against several independent data sets in each species. Although the human model is subject to some uncertainty at higher doses, it has been effectively validated for doses in the range of interest for development of the RfC and RfD. Given the data supporting relatively even distribution between mother and fetus, the extrapolation for exposures *in utero* appears to be well justified. The extrapolation for postnatal exposures, especially lactational, is subject to greater uncertainty, if only because of the diversity of feeding regimes and behaviors in human infants, but as noted in the report this exposure period is probably of less concern than the period *in utero* for sensitivity to the effects *in utero*deemed critical in determining the RfC and RfD. However, the model is justified for calculating exposures in infants with extensive lactational exposure, and these are certainly of concern in terms of health impacts later in the developmental process, such as possible neurobehavioral impacts.

(B) Inhalation Reference Concentration (RfC) for Methanol

B1. A chronic RfC for methanol has been derived from a perinatal inhalation study of the effects from exposing rat dams and pups to methanol during gestation and lactation (NEDO, 1987). Reference values from mouse (Rogers et al., 1993) and monkey (Burbacher et al., 1999; 2004) developmental studies, were also derived and discussed, but were not chosen for the RfC. Please comment on whether the selection of the principal study has been scientifically justified.

The Agency reviewed a substantial number of studies, including those mentioned here, and in those considered adequate from a methodological standpoint dose responses were analyzed for a number of individual data sets. Selection of the principal study was explained and is in accordance with the usual guidelines recommending use of the study with the best data quality (including, in this case, availability of a validated PBPK model) and greatest sensitivity. The F_1 male Sprague-Dawley rats from the study by NEDO (1987) fit these criteria: other analyzable data sets were used as supporting studies. However, the Agency may need to reevaluate the study selection with regard to sensitivity if modification of the PBPK analysis for SD rats significantly alters the relative sensitivity (based on HECs) of the rat, mouse and monkey studies.

B2. Reduction of brain weight at 6 weeks postnatally as reported in the NEDO (1987) developmental rat study was selected as the critical effect. Please comment on whether the rationale for the selection of this critical effect has been scientifically justified. Please identify and provide the rationale for any other endpoints (e.g., other reproductive and developmental effects reported in mouse and monkey studies) that should be considered in the selection of the critical effect.

The report provides a thorough presentation of the various endpoints available for consideration from the NEDO (1987) studies and other reports. The various endpoints are reviewed for sensitivity, statistical significance and general data quality. These criteria and the standard risk assessment procedures, including selection of the most sensitive endpoint (subject to data quality considerations) point to the NEDO (1987) brain weight reduction data as the best choice for the critical effect. The other endpoints provide important

supporting information, but are generally either less sensitive or less reliable (especially the primate studies which necessarily involve fewer maternal subjects per group and smaller litter sizes, as well as showing various deficiencies in the experiments as reported.). However, the Agency may need to reevaluate the endpoint selection with regard to sensitivity if modification of the PBPK analysis for SD rats significantly alters the relative sensitivity (based on HECs) of the rat, mouse and monkey studies. Standard guidelines indicate the choice of the most sensitive strain and endpoint available, subject to limitations of study design and quality.

B3. Benchmark dose modeling of decreased pup brain weight relative to maternal internal methanol doses predicted by the PBPK model was used to derive the point of departure (POD) for the RfC. Has the BMD/PBPK approach been appropriately conducted? Has adequate justification been provided for the selected internal dose metric, i.e., area under the curve (AUC) for methanol, in the blood of dams? Please identify and provide the rationale for any alternative approaches for the determination of the POD, including choice of another dose metric (e.g., methanol metabolized), and discuss whether such approaches are preferred to EPA's approach.

Several possible choices for the dose metric (AUC, peak concentrations, amount of methanol metabolized) are discussed both in the context of their availability from the PBPK model and their possible relevance in the light of studies of the mechanism of action and time-course of exposure during sensitive periods for teratogenesis. Studies of possible modes of action for methanol toxicity have identified the importance of formate in acute toxicity, but imply that this is not a major element of the mechanism for the developmental effects. One possible mechanism for the impacts on the fetus is a direct effect of methanol itself, or of its proximal metabolite formaldehyde. Another possibility is the generation of reactive oxygen species (ROS) as a side effect oxidation of methanol. This hypothesis appears to be popular with some investigators, but actual evidence of its importance seems to be limited. In view of these mechanistic findings the choice of a dose metric based on methanol rather than downstream metabolites appears sensible. The report provides adequate justification for selection of the AUC for methanol as the appropriate internal dose metric in analyzing the NEDO (1987) data, which involve an ongoing exposure time element. (For the single-day exposure experiments by Rogers et al. [1995] the non-time-dependent dose metric C_{max} for methanol was shown to be more suitable, as described in Appendix D.). Even if the metabolism-related formation of ROS or formaldehyde are important contributors to the observed toxic effects, a methanol-based dose metric is applicable when the downstream metabolic processes such as removal of ROS or formaldehyde are much faster than the rate-limiting oxidation of methanol.

Conduct of the BMD analysis was correct and in accordance with the usual guidance for this approach, and was thoroughly reported. The report is correct in noting (on page 5-15) that a 5% BMR is appropriate for analysis of quantal data from the developmental studies considered here: however, it is incorrect to comment as was done here that "a 10% BMR is adequate for moist traditional bioassays". Although this proposal was included in the original guidance for the BMD methodology, subsequent experience by various risk assessors (including Agency staff in a number of recent assessments) have concluded that the 5% BMR is more appropriate for identifying a POD to which the standard UFs are applied (i.e. no UF_L) in a standard animal study with quantal data. Neither the 5% nor the 10% BMR have any particular *a priori* justification for continuous data: the default assumption in this case is the BMR of 1 standard deviation of the control dataset (as preferred here). In any case the data need to be examined to determine an appropriate BMR

representing a minimal detection level or threshold of biologically significant response: this especially applies for continuous data.

B4. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the POD for the derivation of the RfC. It is assumed that these UFs account for variability in methanol dosimetry among human newborns following gestational and lactational exposure, and for uncertainty regarding the ratio of newborn-dose to maternal-dose in humans. Please comment on these assumptions and on the scientific justification for the selected UFs.

The report identifies the potential for significant inter-individual variability in the effects on humans to methanol, including sensitive sub-populations with various enzyme polymorphisms, folate deficiency and enhanced exposure and sensitivity of infants and children. However there does not appear to have been much effort to actually quantify this variability (perhaps due to insufficient data), and the standard default value of UFH = 10 is used. It is a matter of some debate whether this factor is in fact sufficient in the general case, although the ranges of variation in enzyme levels suggested do not appear to be as large as for some other toxicants. Some further examination and discussion of this issue would be helpful in establishing the limits of the data available to inform the decision on the value for UFH.

The use of a UF of 3 to account for the uncertainty in the toxicodynamic component of the extrapolation from rodents to humans is an appropriate application of the default, since no independent data are available to further quantify this extrapolation. Use of a value of 1 for the toxicokinetic component of this uncertainty is also appropriate in view of the use of a comprehensive and validated PBPK model for both humans and the rodent test species. The specific issues of toxicokinetic uncertainty for gestational exposure, and inhalation or lactational exposure of newborns, are discussed in the document.

The question of fetal vs. maternal exposure appears to be relatively well addressed in animal models, and there is no reason to expect major differences for humans since the explanation rests largely on chemistry of the toxicant, rather than the physiology or anatomy of the subject. It therefore appears that this question is not a major source of uncertainty. The exposure and toxicokinetics of the newborn does appear to involve somewhat greater uncertainty, although some of this is associated with differences in diet and behavior and is probably accommodated by the allowance of a UF of 10 for human inter-individual variability. There is no particular reason to expect greater uncertainty in this respect with regard to the interspecies extrapolation.

Use of a factor of 3 for database uncertainty is justified by the absence of rigorous developmental neurotoxicity tests in rodents. Such effects could reasonably be anticipated given the suggestive (but quantitatively inconclusive) results in monkeys, and the clear observation of anatomical impacts on the central nervous system of rodents. This is a source of concern since changes in brain weight imply a relatively substantial change in CNS development which is quite likely to have functional impacts, and in many cases primates have proved to be more susceptible to these effects than rodents. It appears that the results of the developmental study in primates (Burbacher et al, 1999) support this concern. It seems unlikely that this latter study could be used as the critical study for determination of an RfC due to its design limitations, but if it were these limitations would in any case need to be represented by an appropriate UF.

The Agency applied the UFs to the POD expressed as a HEC, which is the standard procedure and is preferred to alternative suggestions that the UFs be applied to intermediate measures such as blood concentrations or AUCs.

(C) Oral Reference Dose (RfD) for Methanol

C1. EPA concluded that the oral RfD should be derived using a route-to-route extrapolation from the more extensive inhalation database given the paucity of oral toxicity data. Please comment on whether the rationale for this approach has been scientifically justified and clearly explained. Please identify and provide the rationale for any alternative approaches for the determining the RfD and discuss whether such approaches are preferred to EPA's approach.

This choice was clearly explained and justified in the report. Given the availability of a detailed PBPK model which has been validated for several species and routes of exposure, it is entirely reasonable to use this model for route-to-route extrapolation from the more extensive and reliable inhalation data to the oral route. Although oral developmental studies are available which suggest similar developmental impacts to those seen by inhalation, these studies used relatively high doses and were therefore not suitable for determining a minimal effect POD on which the RfD could be based. Route-to-route extrapolation is a well established procedure in risk assessment which has been used on a number of previous occasions when appropriate data are available to support it. This is clearly preferable to the alternatives of either developing an RfD based on route-specific but inadequate data, or not developing an oral standard at all.

C2. A PBPK model was used to derive the RfD via a route-to-route extrapolation, in which the internal-dose POD used for the derivation of the RfC based on data from the NEDO (1987) study was extrapolated to human oral exposure levels using the human PBPK model. Please comment on whether the rationale for this approach has been scientifically justified. Has adequate justification been provided for the selected internal dose metric, i.e., AUC for methanol, in the blood of dams? Is the PBPK model suitable for extrapolation of fetal and neonatal endpoints to human oral exposures? Please provide a detailed explanation.

The development of the PBPK model and identification of a POD for the internal dose metric is essentially the same for both inhalation and oral exposures, and has been well described and justified in the report. Similarly, the considerations for selection of the internal dose metric (AUC for methanol) and questions about whether the maternal AUC is reflective of the critical concentrations in the fetus are largely independent of the route. It appears since methanol is readily distributed throughout the aqueous compartments of the body, and that rates of metabolism in the fetus are not so great as to significantly perturb this close-to-equilibrium distribution. Maternal blood concentrations of methanol are therefore expected to be adequate predictors of the concentrations in critical tissues of the fetus for either inhalation or oral exposures. The main difference for the oral route is the discontinuous exposure pattern in rodents (and also humans, although this is harder to define for the general case), and the finite holding time in the stomach prior to absorption into the systemic circulation. The description of the model includes features to address these points, including efforts to validate the model against independent oral data.

Earlier comments (Section A1) as to the possible re-interpretation of model parameters for the SD rat apply for the oral model as well as the inhalation model.

C3. EPA applied the same UFs to the POD for the derivation of the RfD as for the RfC. Please comment on the rationale for the selection of the UFs.

Given the route-specific calculation of oral HED using the PBPK model, and the use of the same underlying toxicity database and critical study, the justification for the UFs is the same as that noted in the discussion of the RfC derivation. There may be some minor differences in the extent of interindividual variation in exposure or susceptibility between routes, but these are probably not sufficient, or sufficiently well characterized, to justify any change in the chosen value of UF_H. It may be argued that the reliance on route-to-route extrapolation involves some additional uncertainty, but given the use of a validated model, and the observation of similar developmental effects in oral studies judged adequate for hazard identification but not for dose-response assessment, there does not seem to be a strong case for changing either UF_A or UF_H on this account. The arguments in favor of a value of 3 for UF_D are exactly the same regardless of route.

(D) General Charge Questions

D1. Is the Toxicological Review logical, clear and concise? Has EPA clearly synthesized the scientific evidence for non-cancer hazards?

The review is thorough and well written, and takes care to provide descriptions of the available evidence in a clear, complete and unbiased form. The report presents a careful and well justified synthesis of these data. The decision to review the toxicokinetics in the first section of the report is a departure from the usual format, but in view of the importance of this topic for methanol and the extent of this section this is a defensible choice.

D2. Please identify any additional studies that would make a significant impact on the conclusions of the Toxicological Review and should be considered in the assessment of the non-cancer health effects of methanol.

No additional studies beyond those presented in the report (and the recent addendum) were identified.

D3. Please discuss research likely to substantially increase confidence in the database for future assessments of methanol.

Although it is unlikely that additional extensive primate studies either could or should be undertaken, it is unfortunate that the main developmental and neurotoxicity studies in non-human primates are deficient in design, for instance using wild-caught animals and lacking a proper control group or dose-response assessment. It would be interesting to see further studies to illuminate the relative sensitivity of rodents and primates to chronic methanol toxicity, especially with regard to developmental and neurotoxicity endpoints. Possibly some further studies *in vitro* would be illuminating.

The lack of an assessment of neurobehavioral impacts on rodent development is a significant data gap, as noted by the Agency in their selection of a database uncertainty factor. Given the reliance of the RfC on an anatomical measure of developmental neurotoxicity (brain weight reduction) which might prove less sensitive than a functional evaluation, this is a significant deficiency which does not seem too difficult to remedy.

(E) Additional Charge Question raised at the Peer Review Meeting

E1. Please comment on the proposed RfD and RfC values for their intended use in risk assessment. Are these numbers more conservative than they need to be to protect public health?

The issue about whether these values are too conservative arises because of the observation that, using the PBPK model to predict the additional levels of blood methanol levels at the RfC or RfD, it appears that these levels result in a relatively small (1 to 15%) increment in blood methanol concentrations relative to the background levels resulting from metabolism and/or dietary exposures.

A point that needs to be emphasized is that the RfC and RfD are specifically defined as levels at which the risk assessor can be reasonably confident that adverse effects will not appear. They are not threshold levels at which effects might start to appear, and there is no generally accepted method, other than the pragmatic consideration of hazard index values, for determining what risks may exist from exposures above the RfC or RfD. However, it is certainly true that this sort of calculation may affect the popular perception of the reasonableness of the calculation, and need to be addressed in discussion of the proposed values.

The proper approach to the question about "normal" background values relative to the RfC and RfD is to ask whether the observed background levels place any constraints on the values of uncertainty factors used in derivation of the protective levels. If a given background level is seen as "safe" (or even beneficial in the case of agents such as fluoride or essential metals), it obviously does not make sense to set values of UF_Aor UF_H which cause the protective levels substantially lower than this level, since this background level can be seen as providing additional evidence on human sensitivity to the agent in question. Unfortunately in the case of methanol this is complicated by the fact that observed background levels in humans appear to show both substantial inter-individual variation and substantial temporal variation in a single individual. Ordinarily, a "safe" level might be identified by considering a point somewhere near the high end of the distribution of background levels. This might be sustainable if you believe that the adverse developmental effects of methanol are strictly caused by methanol itself. However, in view of the uncertainties as to fetal metabolism, mode of action and contribution of diet and individual metabolic or toxicodynamic differences which are identified in the report it seems very unwise to conclude that high-end exposures which are apparently safe for some individuals are *necessarily* safe for all. The assumption that current background levels of methanol are without effect seems plausible, but it is not clear that there have been any analytical investigations of this issue. At the very least an uncertainty factor is needed to reflect these concerns, which therefore indicates that the proposed values for RfC and RfD are not necessarily unreasonable. The Agency needs to address this discussion, which is not covered in the draft report.

EDITORIAL COMMENTS

Page 3-7: The new version of Figure 3-1 in the addendum is a great improvement on the previous figure, which was confusing and uninformative.

Some minor typographic and editorial anomalies were noted while reading the document. These should be rectified in the final version:

Page 3-12: Footnote 9, final sentence: "… this too was not statistically significant." What other non-significant result is referred to here? Clarify, or remove the word "too".

Page 4-4 line 32, page 4-5 line 2, page 4-7 line 1. The antidote/drug fomepizole is referenced on these three occasions, but it is not until page 4-41 line 2 that it is explained that this is chemically 4-methylpyrazole, and that it is an inhibitor of ADH1. It would be useful to have this definition of the chemical name and mode of action pointed out at the first reference rather than the last.

Pages 3-20 to 3-22. The description of the model by Ward et al. (1997) appears twice (as Sections 3.4.2.1 and 3.4.2.3), both appearances being identical except for Table 3-8 which appears only in the second instance. Delete the redundant entry and re-number the sections.

Page 4-1, line 22 refers to "lentiform nuclei". Are these the same as the "lenticular nuclei" referenced on Page 4-2 line 5? If so the terminology should be consistent: if not the difference should be explained.

Page 4-2 lines 18-19. "…Benton and Calhoun (1952) reported on methanol's visual disturbances". What did they say? Also this sentence could be phrased better – it was the 320 individuals who had the visual disturbances, not the methanol.

Page 4-18, line 27. What are "proximal uniferous tubules" in the kidney? Is this a typo for uriniferous? Most people just call them proximal tubules.

Pages 4-30 to 4-31. Table 4-2 would be much easier to read if it did not split over two pages. Some slight abbreviation might be necessary to achieve this.

Page 4-77 line 7. "The data are summarized separate sections…" Need the word "in" after "summarized".

Page B-41 line 33. "did not scaled" should read "scale".

COMMENTS SUBMITTED BY

Lisa M. Sweeney, Ph.D, DABT

Peer Review Meeting of EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Comments Submitted by Dr. Lisa M. Sweeney

(A) Toxicokinetics and PBPK Modeling

A1. Please comment on the scientific soundness of the PBPK model used in this assessment.

Rat Models

The Sprague-Dawley (SD) rat PBPK model is inappropriately parameterized (or insufficiently validated) for the inhalation route. The SD rat model was calibrated based on the iv and oral data. These data have advantages due to the greater certainty in delivered dose as compared to an inhalation study. However, no comparisons of the model predictions to the available inhalation data were made. The LOAEL of the key SD rat study was 1000 ppm. The simulations in Figure B-13 show a predicted blood Cmax of 10.6 mg/L at 1000 ppm (page B-26). The data in Table 3-5 (page 3-5), indicate a post exposure blood concentration of 83 mg/L in SD rats after 8 hrs exposure (Perkins et al., 1995a). (In contrast, on page 3-36 line 24, it is stated that "there are no inhalation data for SD rats".) If that "post exposure" blood concentration was taken immediately at the end of exposure, this datum suggests that the SD rat model is off by a minimum of a factor of 8 at this exposure concentration (if steady state had been achieved). In addition, pages missing from the PDF of the NEDO (1987) report were provided by the Methanol Institute during the review, and indicate thatF1 female SD rats exposed to 1000 ppm methanol had blood concentrations of 99.48 mg/L at the end of exposure. Likewise, the F344 rat model was not tested against the NEDO (2008b) data, where the post exposure blood levels for 1000 ppm exposures again exceed the steady state blood levels depicted in the 1200 ppm model fits (e.g., Figure B-11, page B-22). The conclusions for the F344 rat are of lesser concern for this assessment, since the (current) key toxicology study was conducted using SD rats.

It should be noted that the EPA model uses (for both F344 and SD rats) a fitted fractional inhalation availability (FRACIN) derived for F344 rats (20%) which is much lower than the corresponding values for mice (66.5%) and humans (86.6%). In my use of the model, if FRACIN for SD rats is changed from 0.2 (20%) to 0.6 (60%), the 1000 ppm blood prediction is in agreement with Perkins et al. (1995a), suggesting that the Horton et al. F344 data set is an outlier. When the 60% value is used in the simulation of the kinetics in the NEDO developmental study, the predicted daily AUCs increase by 6.3-, 13.4-, and 22.1-fold at 500, 1000, and 2000 ppm respectively. I did not take this exercise so far as to redo the BMD analysis with the new internal doses, but given that the March 2011 draft BMDL was in the vicinity of the NOAEL (500 ppm), it seems reasonable to expect that a revised BMDL (and candidate RfC) would also increase by ~6 fold.

EPA does not provide a sensitivity analysis of the rat PBPK model, even though PBPK-derived dose metrics from a rat study provide the basis of the RfC and RfD. A sensitivity analysis of the rat blood methanol AUC under conditions approximating the BMDL would appropriately focus the evaluation of model reliability on key model parameters.

Human Model

I have concerns regarding the parameterization/validation of the human PBPK model and the lack of human model sensitivity analyses, factors which undermine the confidence in the model and its application. Useful human kinetic studies were apparently overlooked by EPA and their contractors (see Section D2, below). While these studies are limited in the number of subjects involved, they are potentially quite valuable in model parameterization because they do not involve the inhalation route. The inhalation studies can have some substantial dosimetric uncertainties due to the possible interspecies differences in fractional uptake and uncertainty with respect to breathing rates. The human iv data (Haffner et al., 1992) and oral data (Schmutte et al., 1988) are not subject to the same dosimetric uncertainty.

I also do not understand why EPA did not report any results from human model sensitivity analyses. Ideally, EPA would have conducted sensitivity analysis on steady-state blood methanol concentrations at the HEC and HED or RfC and RfD or values in between (see below, section A.2.). These analyses would focus the model confidence assessment on the parameters that are the key determinants of the internal dose and would inform the choice of UFH.

At a minimum, EPA should assess whether or not the model they used in the risk assessment can (adequately) simulate the additional human data identified herein and conduct and provide human model sensitivity analyses at the RfC and RfD. EPA should further consider reparameterizing the human methanol PBPK model.

Mouse Model

It seems odd that, for oral dosing, the mouse blood levels are reported to be insensitive to any parameter related to clearance (e.g., metabolism, blood flow to the liver) (pp B-16 and B-18). It is also not clear what type of oral dose is being simulated based on the text alone (appears to be gavage in model files on-line). The runtime files that should reproduce Figures B-2 and B-5 yield simulations that are slightly off.

EPA does not provide files that fully recreate the sensitivity analyses--only those parameters demonstrated in Figures B-6, B-7, and B-8. These files do not accurately reproduce the figures in the document. I tested an additional parameter that does not appear on the figures (FRACIN), and found that the model output was, as I expected, very sensitive to this parameter. Thus the sensitivity analysis does not appear to have been comprehensive.

A2. Please comment on the scientific justification for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks.

The subtraction of background levels (or the alternative modeling with background included) does not appear to have made a significant difference on the PBPK modeling of the available kinetic studies or BMD analysis.

Having said that, it is my opinion, that if the object is to understand risk, and risk is related to internal dose, what we should be doing is attempting to define an acceptable (human) internal dose, and then, from that, derive acceptable external exposures. The sticking point with the methanol assessment is the interpretation of background levels and their contribution to risk. A fundamental difficulty, then, with this noncancer assessment, is that the internal dose BMDL (internal dose point of departure, iPOD) is converted to an HEC/HEC, and the uncertainty factors are applied to the HED/HEC to derive the RfD/RfC. Then, when the RfC/RfD is simulated, it is then realized that the additional methanol body burden is indistinguishable from background.

A more direct route to the comparison to background is to take the iPOD, convert it to a time-weighted average (TWA) concentration (daily AUC divided by 24 hrs) and divide by the UFs directly (90.86 hr x mg/L/24 hrs = 3.8 mg/L; 3.8 mg/100 = 0.038 mg/L). It is immediately obvious that a difference of 0.038 mg/L methanol against a background of ~1.8 mg/L (with standard deviations of about 0.7 to 1.2 mg/L) is miniscule and thus is a de minimis increase in population risk. If you accept the UFA of 3, UFD of 3, and the pharmacodynamic component of UFH as 3, (I do not—see B.4), this gives you a composite of UF of 30 prior to consideration of human TK variability. If the acceptable additional blood methanol concentration above background is thus defined, with 3.8 mg/L ÷ 30 = 0.13 mg/L, the RfC would be ~5 ppm (see page B-37). I would like to see a sensitivity analysis on the *sum* of background methanol (~1.8 mg/L) plus additional methanol (0.13 mg/L) at 5 ppm. I doubt that total blood methanol would be very sensitive to any anticipated biological variability or parameter uncertainty under these circumstances. Thus the TK component of UFH can reasonably be 1. Even with a composite UF of 10 (e.g., with UFD = 1), the iPOD/UFc = 0.38, which is less than half the reported standard deviation of background methanol in most of the studies noted by EPA.

A3. The PBPK modeling effort assumed similar methanol pharmacokinetics between pregnant and non-pregnant animals. Please comment on the adequacy of the dose-metric extrapolation based on a PBPK model for non-pregnant adults (i.e., no fetal compartment) for predicting risks associated with fetal/neonatal brain concentrations of methanol.

The extrapolation is adequate.

A4. EPA assumes limited methanol metabolism in the fetus because of limited alcohol dehydrogenase (ADH) activity in the human fetus, limited catalase and ADH activity in fetal rodents, and existing pharmacokinetic data that show nearly equal concentrations in maternal blood vs. the fetal compartment. Please comment on the validity of this assumption given the lack of data regarding potential alternate metabolic pathways in the fetus.

The assumption appears valid.

A5. Please comment on the scientific justification of the extrapolation approach from rats to humans for in-utero and neonatal lactational and inhalation exposures.

Conceptually, the rat POD extrapolation to the human HED and HEC appear to be acceptable; the rat POD in and of itself, however, appears questionable. The human model parameterization also needs to be further evaluated (see A1).

(B) Inhalation Reference Concentration (RfC) for Methanol

B1. A chronic RfC for methanol has been derived from a perinatal inhalation study of the effects from exposing rat dams and pups to methanol during gestation and lactation (NEDO, 1987). Reference values from mouse (Rogers et al., 1993) and monkey (Burbacher et al., 1999; 2004) developmental studies, were also derived and discussed, but were not chosen for the RfC. Please comment on whether the selection of the principal study has been scientifically justified.

The selection of the principal study is contingent upon the details of the dose-response analysis and determination of the HEC/HED. Comparison of the SD rat PBPK model output for inhalation exposure with the experimental data suggests that the rat POD could be on the order of 6-fold too low (see A1). A revised rat POD would be similar to the mouse values. Regarding the monkey study, I do not believe it provides convincing evidence of an effect, given the inconsistencies in dose-response, multiple comparisons, and the potential for unreliable identification of "effects" in small studies. Admittedly, these areas are generally outside of my particular expertise.

B2. Reduction of brain weight at 6 weeks postnatally as reported in the NEDO (1987) developmental rat study was selected as the critical effect. Please comment on whether the rationale for the selection of this critical effect has been scientifically justified. Please identify and provide the rationale for any other endpoints (e.g., other reproductive and developmental effects reported in mouse and monkey studies) that should be considered in the selection of the critical effect.

As noted for B1, the discrepancies between the SD rat inhalation PBPK model and data lead me to question the internal dosimetry estimates that led to this study being identified as "critical".

B3. Benchmark dose modeling of decreased pup brain weight relative to maternal internal methanol doses predicted by the PBPK model was used to derive the point of departure (POD) for the RfC. Has the BMD/PBPK approach been appropriately conducted? Has adequate justification been provided for the selected internal dose metric, i.e., area under the curve (AUC) for methanol, in the blood of dams? Please identify and provide the rationale for any alternative approaches for the determination of the POD, including choice of another dose metric (e.g., methanol metabolized), and discuss whether such approaches are preferred to EPA's approach.

The dose metric (blood methanol) is appropriate, but does not appear to have been reliably computed for rats. The BMD approach appears to have otherwise been appropriately implemented.

B4. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the POD for the derivation of the RfC. It is assumed that these UFs account for variability in methanol dosimetry among human newborns following gestational and lactational exposure, and for uncertainty regarding the ratio of newborn-dose to maternal-dose in humans. Please comment on these assumptions and on the scientific justification for the selected UFs.

The uncertainty factors UFH and UFD are inadequately justified.

At the level of the proposed RfC and RfD, intraspecies differences in disposition of exogenous methanol in humans will likely have no meaningful impact on the body burden of "total" methanol. Thus a full UFH of 10 is not warranted. Sensitivity analyses of the human PBPK model for methanol could identify parameters which have an impact on total or "additional" methanol, but were not conducted. Identification of the potential impact of variability and/or uncertainty of the human model parameters on predicted body burden of methanol could further inform the selection of UFH values. The discussion of UFH inappropriately includes the special sensitivity of children. Since the database includes two-generation studies (in fact, the current key study is a two-generation study), there is no reason discuss children's potential susceptibility; no particular developmental susceptibility of humans vs. test species is expected. The authors appear to be attempting to double dip on an uncertainty factor that is not needed.

It is hard to imagine how a UFD of 3 can be justified. As the authors note, the key endpoint is developmental toxicity, which has been evaluated in multiple species, including primate, and special endpoints such as neurotoxicity and immunotoxicity have been evaluated. There is no need to have a UF because "there is uncertainty regarding which test species is most relevant to humans"—the lowest, high-quality point of departure was used. There is also no need to have a UFD for "dose spacing" because the BMD analysis counters this potential design deficiency.

Also, I personally believe it is most appropriate to apply the uncertainty factors to the internal dose point of departure, prior to interspecies extrapolation with the pharmacokinetic model to account for non-linearities in external vs. internal dose relationships. EPA should discuss their choice of applying UFs to the HED rather than the BMDL. While this point is moot when the kinetics are linear with respect to exposure concentration, this is not the case with methanol. In the case of EPA's March 2011 analysis for the rat, the difference appears to be small (RfC of 1.8 mg/m^3 using HEC/UF, 2.0 mg/m^3 using BMDL/UF, based on my calculations, ~10%), but the difference becomes larger starting from a higher point of departure. For example, if UFs are applied first to the mouse cervical rib BMDL05, then converted to the candidate RfC using the PBPK model, the candidate RfC increases from 10.4 mg/m^3 to 23.6 mg/m^3 (>2-fold).

(C) Oral Reference Dose (RfD) for Methanol

C1. EPA concluded that the oral RfD should be derived using a route-to-route extrapolation from the more extensive inhalation database given the paucity of oral toxicity data. Please comment on whether the rationale for this approach has been scientifically justified and clearly explained. Please identify and provide the rationale for any alternative approaches for the determining the RfD and discuss whether such approaches are preferred to EPA's approach.

Route-to-route extrapolation appears to be justified. Human model validation using the oral data of Schmutte et al. (1988) (see D.2) would further strengthen confidence in the route-to-route extrapolation.

C2. A PBPK model was used to derive the RfD via a route-to-route extrapolation, in which the internal-dose POD used for the derivation of the RfC based on data from the NEDO (1987) study was extrapolated to human oral exposure levels using the human PBPK model. Please comment on whether the rationale for this approach has been scientifically justified. Has adequate justification been provided for the selected internal dose metric, i.e., AUC for methanol, in the blood of dams? Is the PBPK model suitable for extrapolation of fetal and neonatal endpoints to human oral exposures? Please provide a detailed explanation.

With the exception of numerical value of the point of departure and the order in which UFs are applied/extrapolations made, the approach appears to generally be valid. Maternal blood is an adequate surrogate for fetal dosimetry, based on the available data.

C3. EPA applied the same UFs to the POD for the derivation of the RfD as for the RfC. Please comment on the rationale for the selection of the UFs.

My comments above (question B4) apply equally to the UFs for the RfD.

(D) General Charge Questions

D1. Is the Toxicological Review logical, clear and concise? Has EPA clearly synthesized the scientific evidence for non-cancer hazards?

<u>Major comments</u>

EPA has inadequately synthesized the SD rat toxicokinetic data in their PBPK model (see A.1).

While chapter 3 (Toxicokinetics) and Appendix B (PBPK model) cover many of the same data sets, it is clear that they were written by different people and that no one took the time to ensure that the reader can seamlessly move between these sections, as one would expect in a more cohesive document. For example, key studies from the PBPK modeling are omitted from Chapter 3 summary tables 3-2, 3-4, and 3-5 (Ernstgard et al., 2005; Rogers et al., 1997, Ward et al., 1997). While the same data may have been presented in the studies that are listed, it is much harder to follow as is. Also, relevant studies noted in chapter 3 were not used in model development/validation (see A1).

The clarity of the document is hampered by the lack of a clear synthesis of evidence regarding plausible modes of action for developmental toxicity.

Minor comments

Page 3-3. This table (from CERHR 2004) is not up to date, omitting the data of Ernstgard et al., 2005.

Page 3-5. The reported blood methanol levels for the NEDO studies are unclear. Are the "0" ppm data the background methanol levels, and the 10, 100, and 1000 ppm entries measured values minus background?

Page 3-13. Methanol blood data for monkeys are discussed, but not shown, while formate and folate data are shown. Considering EPA proposes that the MOA and key dose metric are related to methanol, not the metabolites, it seems that the methanol data should be shown, and the formate and folate data are less important.

Page 3-25, Table 3-9. This table is not as informative or useful as it could be. Key studies are omitted (e.g., Ernstgard et al., 2005). Organizing the data by alphabetizing the first authors is less helpful than grouping by species and strain. Some of these data were not used for the EPA-developed PBPK model (e.g., Gonzalez-Quevedo et al., 2002). In some places, animals are described only as pregnant, while in others, Gestation Days are clearly specified. It is not clear where duplicate reporting of same data has occurred (e.g., Perkins et al., 1995a, 1995b, 1996 and Pollack and Brouwer, 1996, Pollack et al. 1993).

Page 5-11, Table 5-2. Clarity would be improved by converting steady state daily AUC to time-weighted average concentration. These concentrations could more readily be compared to the experimental data and are more readily understood by non-experts.

Page 5-14, Figure 5-1. The figure caption should provide the units for the response and dose; better yet, this information should be on the figure's axes. This comment applies equally to all of Appendix C.

Page B-8. The metabolic parameters for the mouse do not correspond to "one high affinity/low capacity and one low-affinity, high-capacity enzyme". The lower affinity enzyme, as described by Vmaxc2 and KM2, is also much lower capacity (3.2 vs. 19).

Page B-38. Daily AUC can and should be converted to TWA concentration as noted above (comment on page 5-11).

Page B-40. At this point, the authors introduced the reference Pollack and Brouwer, 1996, which appears to cover data published elsewhere (e.g., Perkins et al.). This was initially confusing, because I thought new data were being introduced. (I read Appendix B prior to reading Chapter 3.)

Page C-42, line 11. It is incorrectly stated that the monkey BMD modeling was done on the basis of external concentration (ppm).

Page C-44, Table C-10, footnote a. It is incorrectly stated that the AUC for the monkey study was estimated using a rat PBPK model.

Nit-picky comments (e.g., typographical errors, unclear referencing.)

Pages 3-20, 3-21. The text of section 3.4.2.1 is identical to the text of section 3.4.2.3.

Page 4-6, line 3. "improved" should be "improve".

Page 5-4, line 25. I don't think young monkeys are called "pups".

Page 5-6. Footnote "a" is not found in the table.

Page B-5, Line 3. PPK should be PBPK.

Page B-6. Table is hard to read (biochemical constants squished).

Page B-7. Footnote g was not found in the table. Footnote k was not consistent with the text (i.e., source in table says Ernstgard et al., 2005, but Sedivec et al. (1981) in footnote).

Page B-10, Line 26. The text about KMASC is not consistent with Table B-1.

Page B-11. Listing one data source within the figure caption and one below the figure caption is confusing.

Page B-13, Fig B-5. Legend says GD8, caption says GD9. Text (p B-14) says GD9. Text mentions 15 hr data, but none is evident in the figure.

Page B-20, line 6. "know" should be "known".

Page B-27, lines 1 and 2. This sentence needs a verb.

Page B-27 line 6. "serious" should be "series".

Page B-28. Daily dose should be in figure caption.

Page B-35, lines 9 and 10, figure caption. Exponents should be properly formatted. Remove extra period in source citation.

Page B-41, line 33. "scaled" should be "scale".

D2. Please identify any additional studies that would make a significant impact on the conclusions of the Toxicological Review and should be considered in the assessment of the non-cancer health effects of methanol.

Haffner HT, Wehner HD, Scheytt KD, Besserer K. The elimination kinetics of methanol and the influence of ethanol. Int J Legal Med. 1992; 105(2):111-4. This study presents blood concentration time course data for methanol infused iv in a healthy male volunteer. Data for 3 other individuals were not shown, but indicated similar blood half lives.

Schmutte P, Bilzer N, Penners BM. Zur Nűchternkinetik der Beglietalkohole Methanol und Propanol-1. [Kinetics of the congeners methanol and propanol-1 in the absence of ethanol]. Blutalkohol. 1988; 25(3):137-42. This paper (in German) presents the blood methanol time course in 5 or 7 volunteers who ingested a methanol/drinking water mixture over a period of 15 minutes.

Incorporation of human kinetic studies by the iv and oral routes in PBPK model development has the potential to change the estimates of the chemical-specific parameters in the PBPK model, and thus the HEC, HED, RfC and RfD.

D3. Please discuss research likely to substantially increase confidence in the database for *future* assessments of methanol.

Inhalation kinetic data for Sprague-Dawley rats appear to be limited.

Monkey studies with longer exposure durations and similar endpoints could be informative.

Additional mode-of-action motivated studies would be helpful.

"Bonus" Question

Please comment on the proposed RfD and RfC values for their intended use in risk assessment. Are these numbers more conservative than they need to be to protect public health?

In a word: Yes.

It does not seem logical to conclude that methanol exposures that increase human blood concentrations from ~1.8 mg/L to ~1.84 mg/L constitute a threshold for meaningful increases human risk. (See A.2.)

How did we get here? The answer seems to be: the combination of excessive UFs with a point of departure that is lower than it should have been (due to an incomplete consideration of the rat inhalation toxicokinetic database). The proposed UFs are excessive in light of the available toxicological database and minimal contribution of additional methanol to the body burden at concentrations/doses relevant to the RfC and RfD

(see B4). More appropriate choices (higher point of departure, lower UFs) may yield RfC and RfD values that are less of a departure from common sense, yet still provide adequate health protection.

The BMDLs were derived on the basis that changes in brain weight should not exceed 1 standard deviation of the natural background variation. It is not unreasonable to extend that reasoning to increases in a naturally occurring background chemical measured in blood to provide a reality check on the reference values derived by the POD/UF approach.

Appendix B: List of Reviewers

United States
Environmental Protection Agency
National Center for Environmental Assessment

Peer Review Workshop for EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Hilton Raleigh Durham Airport
Durham, NC
July 22, 2011

Peer Reviewers

Janusz Z. Byczkowski, Ph.D.
Independent Consultant
Fairborn, OH

Thomas M. Burbacher, Ph.D.
Professor
University of Washington
Seattle, WA

David C. Dorman, Ph.D.
Professor
NCSU-College of Veterinary Medicine
Raleigh, NC

Kenneth McMartin, Ph.D.
Professor
LSU Health Sciences Center
Shreveport, LA

Stephen Roberts, Ph.D. (Chair)
Professor
University of Florida
Gainesville, FL

Andrew Salmon, Ph.D.
Senior Toxicologist
California EPA- OEHHA
Lafayette, CA

Lisa M. Sweeney, Ph.D.
Consultant
Henry M. Jackson Foundation for the Advancement of Military Medicine
Naval Medical Research Unit-Dayton
Kettering, OH

Appendix C: List of Observers

United States
Environmental Protection Agency
National Center for Environmental Assessment

Peer Review Workshop for EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Hilton Raleigh Durham Airport
Durham, NC
July 22, 2011

Final Observers List

Nancy Beck *(via teleconference)*
Office of Management and Budget
Washington, DC

Norman Birchfield *(via teleconference)*
U.S. Environmental Protection Agency
Washington, DC

Vincent Cogliano
U.S. Environmental Protection Agency
Washington, DC

Jan Connery *(Facilitator)*
ERG
Lexington, MA

George Cruzan
ToxWorks
Bridgeton, NJ

Allen Davis
U.S. Environmental Protection Agency
Research Triangle Park, NC

Bridget DiCosmo *(via teleconference)*
Inside EPA
Arlington, VA

Greg Dolan
Methanol Institute
Washington, DC

Charles Elkins
Chuck Elkins and Associates
Washington, DC

Jeff Gift
U.S. Environmental Protection Agency
Research Triangle Park, NC

Patricia Gillespie
U.S. Environmental Protection Agency
Research T riangle Park, NC

Laurie Holmes *(via teleconference)*
American Forest & Paper Association
Washington, DC

William Kelce
Exponent
Cary, NC

Katharine Kurtz *(via teleconference)*
Navy and Marine Corps Public Health Center
Portsmouth, VA

Teresa Leavens
North Carolina State University
Cary, NC

Geniece Lehmann *(via teleconference)*
U.S. Environmental Protection Agency
Research Triangle Park, NC

Lisa Lowe
U.S. Environmental Protection Agency
Research Triangle Park, NC

Resha Putzrath *(via teleconference)*
Navy and Marine Corps Public Health Center
Portsmouth, VA

John Rogers
U.S. Environmental Protection Agency
Research Triangle Park, NC

Reeder Sams
U.S. Environmental Protection Agency
Research Triangle Park, NC

Paul Schlosser
U.S. Environmental Protection Agency
ORD, NCEA
Research Triangle Park, NC

Thomas Starr
TBS Associates
Raleigh, NC

John Vandenberg
U.S. Environmental Protection Agency
Research Triangle Park, NC

Debra Walsh
U.S. Environmental Protection Agency
Research Triangle Park, NC

Appendix D: Agenda

Peer Review Meeting for EPA's Draft Toxicological Review of Methanol (Non-Cancer)

Hilton Raleigh Durham Airport Hotel
Durham, NC
July 22, 2011

Agenda

8:00 a.m. Registration/check in

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

8:45 a.m. **EPA Welcome Remarks** ..... *Vincent Cogliano, Director, EPA IRIS Program*

8:55 a.m. **Public Comment** .....*Jan Connery*

9:20 a.m. **Introduction to the Draft Toxicological Review**.....*Jeff Gift, EPA NCEA*

9:40 p.m. **(D)General Charge Question D1**.....*Stephen Roberts (Chair) & Panel*

D1. Is the Toxicological Review logical, clear and concise? Has EPA clearly synthesized the scientific evidence for non-cancer hazards?

10:15 a.m. **(A) Toxicokinetics and PBPK Modeling** .....*Stephen Roberts* (*Chair) & Panel*

A1. Please comment on the scientific soundness of the PBPK model used in this assessment.

10:50 a.m. BREAK

11:05 a.m. **(A) Toxicokinetics and PBPK Modeling (cont.)** .....*Stephen Roberts* (*Chair) & Panel*

A2. Please comment on the scientific justification for the subtraction of background levels of methanol from the data in relation to the quantification of non-cancer risks.

A3. The PBPK modeling effort assumed similar methanol pharmacokinetics between pregnant and non-pregnant animals. Please comment on the adequacy of the dose-metric extrapolation based on a PBPK model for non-pregnant adults (i.e., no fetal compartment) for predicting risks associated with fetal/neonatal brain concentrations of methanol.

A4. EPA assumes limited methanol metabolism in the fetus because of limited alcohol dehydrogenase (ADH) activity in the human fetus, limited catalase and ADH activity in fetal rodents, and existing pharmacokinetic data that show nearly equal concentrations in maternal blood vs. the fetal compartment. Please comment on the validity of this assumption given the lack of data regarding potential alternate metabolic pathways in the fetus.

A5. Please comment on the scientific justification of the extrapolation approach from rats to humans for in-utero and neonatal lactational and inhalation exposures.

12:20 p.m. LUNCH

Agenda (cont.)

1:20 p.m. **(B) Inhalation Reference Concentration (RfC) for Methanol** .....*Stephen Roberts (Chair) & Panel*

B1. A chronic RfC for methanol has been derived from a perinatal inhalation study of the effects from exposing rat dams and pups to methanol during gestation and lactation (NEDO, 1987). Reference values from mouse (Rogers et al., 1993) and monkey (Burbacher et al., 1999; 2004) developmental studies, were also derived and discussed, but were not chosen for the RfC. Please comment on whether the selection of the principal study has been scientifically justified.

B2. Reduction of brain weight at 6 weeks postnatally as reported in the NEDO (1987) developmental rat study was selected as the critical effect. Please comment on whether the rationale for the selection of this critical effect has been scientifically justified. Please identify and provide the rationale for any other endpoints (e.g., other reproductive and developmental effects reported in mouse and monkey studies) that should be considered in the selection of the critical effect.

B3. Benchmark dose modeling of decreased pup brain weight relative to maternal internal methanol doses predicted by the PBPK model was used to derive the point of departure (POD) for the RfC. Has the BMD/PBPK approach been appropriately conducted? Has adequate justification been provided for the selected internal dose metric, i.e., area under the curve (AUC) for methanol, in the blood of dams? Please identify and provide the rationale for any alternative approaches for the determination of the POD, including choice of another dose metric (e.g., methanol metabolized), and discuss whether such approaches are preferred to EPA's approach.

B4. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the POD for the derivation of the RfC. It is assumed that these UFs account for variability in methanol dosimetry among human newborns following gestational and lactational exposure, and for uncertainty regarding the ratio of newborn-dose to maternal-dose in humans. Please comment on these assumptions and on the scientific justification for the selected UFs.

2:45 p.m. BREAK

3:00 p.m. **(C) Oral Reference Dose (RfD) for Methanol** .....*Stephen Roberts (Chair) & Panel*

C1. EPA concluded that the oral RfD should be derived using a route-to-route extrapolation from the more extensive inhalation database given the paucity of oral toxicity data. Please comment on whether the rationale for this approach has been scientifically justified and clearly explained. Please identify and provide the rationale for any alternative approaches for the determining the RfD and discuss whether such approaches are preferred to EPA's approach.

C2. A PBPK model was used to derive the RfD via a route-to-route extrapolation, in which the internal-dose POD used for the derivation of the RfC based on data from the NEDO (1987) study was extrapolated to human oral exposure levels using the human PBPK model. Please comment on whether the rationale for this approach has been scientifically justified. Has adequate justification been provided for the selected internal dose metric, i.e., AUC for methanol, in the blood of dams? Is the PBPK model suitable for extrapolation of fetal and neonatal endpoints to human oral exposures? Please provide a detailed explanation.

C3. EPA applied the same UFs to the POD for the derivation of the RfD as for the RfC. Please comment on the rationale for the selection of the UFs.

4:00 p.m. **(D) General Charge Questions D2 and D3**.....*Stephen Roberts (Chair) & Panel*

D2. Please identify any additional studies that would make a significant impact on the conclusions of the Toxicological Review and should be considered in the assessment of the non-cancer health effects of methanol.

D3. Please discuss research likely to substantially increase confidence in the database for *future* assessments of methanol.

4:30 p.m. **Reviewer Final Comments**.....*Stephen Roberts (Chair) & Panel*

4:55 p.m. **Closing Remarks**.....*Jan Connery*

5:00 p.m. ADJOURN