PRESENTATION OF DISSENTING OPINION AND QUESTIONS FOR THE AUGUST 17-18 SAP MEETING ON MALATHION

Brian Dementi, Ph.D., DABT, Toxicologist, Health Effects Division, OPP
July 26, 2000

Concerning the assessment of the carcinogenicity of malathion as performed by the Health Effects Division's Cancer Assessment Review Committee (CARC), my views have been documented in the various reviews and attachments to the April 28, 2000 CARC report. Given the large amount of information and the complexity of the subject, I am only able to summarize the information at this point. I pray that SAP members will critically evaluate the background materials.

Three of the attachments are focal documents. The first two of these set forth principles which I often refer to as the bases for my interpretation of data sets containing evidence of neoplasia. ***Therefore, I would be particularly interested to have the SAP comment on the utility of the principles referred to in these attachments, and likewise whether I have been consistent and reasonable in their application***. The third provides an overview of my evaluations of the various neoplastic end points, and is the document after which the current memorandum is patterned and updated. The attachments in question are:

1) The letter of November 12, 1999 to William Burnam, CARC Chairman (Attachment 18). Particularly noteworthy in this letter is how evidence of carcinogenicity occurring at excessive doses should be interpreted. One source referenced says: "The EMTD (estimated maximum tolerated dose) is determined on the basis of prechronic tests and other relevant information. If the test reveals that the EMTD is too high to meet the conditions defined herein, *positive results* (emphasis added) obtained *above the EMTD* (emphasis added) are acceptable as evidence of carcinogenicity unless there is convincing evidence to the contrary. Alternatively, *negative results* (emphasis added) obtained *above the EMTD* (emphasis added) are considered inadequate unless particularly strong and specific scientific reasons justify their acceptance as negative. *Positive results* (emphasis added) obtained at or below the EMTD provide evidence of carcinogenicity." [Interagency Regulatory Liaison Group (IRLG) (1979), *JNCI*, 63, p. 251] Also, further along in the letter: "A *negative* (emphasis added) study is ordinarily accepted by regulatory agencies if:.....survival of all groups (per sex per dose) is no less than 50%.... at 104 weeks for rats." [White House Office of Science and Technology Policy (OSTP)(1985), Fed. Reg. Vol. 50, p. 10414] There is more authoritative information set forth in this attachment that serves to support the reasonableness of the particular passages I have elected to present here. The text cited above is important because interpretations must be made on findings in malathion studies at doses which CARC has concluded are excessive.

2) The letter of June 7, 1999 to William Burnam, CARC Chairman (Attachment 11) This letter, as an example, references a definition of carcinogen that I have cited in many other documents in this package. The statement is quoted from the same OSTP (1985) document as referenced above, and reads as follows: *"A chemical carcinogen may be a substance which either significantly increases the incidence of cancer in animals or humans or significantly decreases the time it takes a naturally occurring (spontaneous) tumor to develop relative to an appropriate background or control group. Either phenomenon is said to represent the effects of a carcinogen."* (pp. 10414-15). In interpreting the tumor data I have attempted to present information that not only addresses incidence data (consistent with the first aspect of the effects of a carcinogen) but evidence that supports enhanced development of a tumorigenic response (consistent with the second aspect of the effects of a carcinogen). I feel CARC has generally placed too much emphasis upon strict statistical analysis of incidence data to the exclusion of evidence of enhanced development of a neoplastic response. Along these lines, I am particularly concerned that in various interpretations, elements of the second aspect be properly recognized in the weight of the evidence when incidence data may not quite achieve statistical significance by the $p = 0.05$ criterion of significance. Again, I have made note of this in several interpretations.

3) The letters of January 27, 2000 and February 3, 2000 (Attachments 27 and 28) to John Carley, Office of Pesticide Programs

These letters consolidate my differences of opinion and conclusions with respect to those of the cancer committee, and serve as a guide for this present summary.

I have engaged in a good deal of research in attempting to interpret the data base. This has involved reading many publications and consulting with several external experts, most notably from the fields of pathology and statistics. I have documented many of these references and results of conversations in my various reviews and memoranda to the CARC. I feel as though the conclusions I have drawn are therefore not simply of my own creation, but in good measure carry the weight of outside expertise and opinion. In fact, I would acknowledge at this point the contributions of those experts who have so willingly been of assistance. I might also add the assistance of persons in NTEU, and also from some other rather surprising sources.

Before proceeding with a summary of assessments of the various neoplastic findings, I should note three principle areas of concern that I would hope SAP members will attend to as they evaluate the materials:

- the evidence of neoplasia in the lowest dose range, 0-500 ppm, and in certain cases extending even to the lowest dose of about 100 ppm.

- the evidence of several neoplastic findings in the various tissues as identified from all

studies in the data base and the collective weight of these in the classification of the chemical as to carcinogenicity.

- whether there has been bias toward removal or marginalization of positive neoplastic findings, manifested in the form of a) re-evaluations *only* of several initially positive neoplastic findings, b) inadequate follow up in obtaining critical information prior to reaching conclusions and c) discounting by whatever rationale the tumorigenic findings that remain after the various pathology re-evaluations.

I. Liver tumorigenic response in the B6C3F1 mouse

In my view, this study should be accepted as providing positive evidence of carcinogenicity. Hepatocellular tumor incidences as reported in the original study submission and as revised by PWG are presented in the CARC report (Tables 1 and 2, pp 6-7). It is clear there were positive findings of carcinogenicity in both sexes at both of the high test dose levels (8,000 and 16,000 ppm). The PWG was performed only on the male histopathology. In the original study submission, statistically significant increases in combined tumor incidence were seen at the 100, 8,000 and 16,000 ppm dose levels, but not at the 800 ppm dose level, though there was a greater than four-fold numerical increase of tumor incidence at that level versus the control. The tumorigenic response in males was characterized by a highly significant trend test ($p = 0.000$). Following the PWG, the male hepatocellular tumor incidences remained highly significantly increased at the 8,000 and 16,000 ppm levels, but no longer so at 100 ppm ($p = 0.075$), nor was it significant at 800 ppm. The trend test remained highly positive ($p = 0.000$). A principle reason accounting for the lack of significance (at the $p = 0.05$ level) for the 100 ppm group following the PWG was the increased incidence of adenomas in the control, from 1/54 (2%) to 4/54 (7%), while the actual number of mice affected in the 100 ppm group remained unchanged at 10/54 (19%).

The difficulty I have with CARC's interpretation of the data (even after the PWG), as expressed in several memoranda [Review of PWG (Doc No. 013721; MRID 44554901); Attachments 1 (11/26/97), 4 (5/4/98), 5 (5/29/98) 12 (6/21/99), 18 (11/12/99)], and in my comments on various draft reports of CARC meetings, is summarized as follows:

1) The defacto discounting for risk assessment purposes of the remarkable tumorigenic responses in both sexes at the top two dose levels, based upon an argument of excessive dosing as evidenced by pronounced inhibition of the blood borne cholinesterases (p. 12 of CARC report) in the absence of cholinergic clinical signs or increased mortality. This is, in my opinion, without scientific justification, particularly at 8000 ppm where brain cholinesterase inhibition was but 23% (males), 20% (females), not statistically significant, and very close to the limit of detection of cholinesterase inhibition under current methodology. It cannot be said for certainty that brain cholinesterase was even inhibited in either sex at 8000 ppm,

but if so, it was marginal. In my witness brain cholinesterase inhibition of 37-43% as recorded for the 16000 ppm group does not qualify as severe. In considering dosing as excessive, one should also observe that in this same study at *800 ppm*, plasma and red cell cholinesterases in females were significantly inhibited by 36% and 58%, respectively. One might ask, in CARC's experience what degree of cholinesterase inhibition qualifies as "severe"? In other words, SAP should be advised what level, or per cent inhibition, OPP considers excessive (with rationale) and was used as a guideline in making this decision [note by contrast in the malaoxon study among females at 1000 ppm there was "severe inhibition of cholinesterase" (p. 35 of the CARC report), and, in contrast to the mouse study, was *accompanied by statistically significant increased mortality (44% versus 13% in control)*, yet this was considered by CARC as an acceptable dose level]. So, I disagree that cholinesterase inhibition (particularly the blood borne enzymes) should be used to conclude dosing was excessive, absent any guidelines or precedent I know of for this conclusion. Please see Attachments 1 (11/26/97) and 12 (6/21/99) for further rationale.

2) The Agency had required the registrant conduct the new mouse study, at the same dose levels of 8,000 ppm and 16,000 ppm as were employed in an older (1978) NCI study in which there was an equivocal increase in hepatocellular tumors in males only, and only at 16,000 ppm. Yet, the new 18-month carcinogenicity study exhibited the remarkable findings in both sexes at both of these dose levels. So the study answered the Agency's question as to the reality of the tumorigenicity at 8,000 and 16,000 ppm clearly in the affirmative, only to have the findings now discounted on the grounds dosing was excessive, based upon cholinesterase inhibition. I do not accept this rationale as adequate to disregard use of this study in quantitative risk assessment for carcinogenicity. Earlier, during the course of consideration of the data base, CARC calculated a Q* for carcinogenicity based upon the tumor incidences for females in this study, only to drop this from its final report, proceeding without any quantitative risk assessment.

3) Not only do I disagree with CARC's deletion of the quantitative assessment of carcinogenicity, but consider as particularly persuasive the evidence of a positive tumorigenic response at the lowest dose level, 100 ppm, in male mice. The reasons for this include:

   a) Even post-PWG, the pairwise comparison ($p = 0.075$) of this group versus control, taken in concert with the remarkable positive trend ($p = 0.000$) should not be discounted as real even in the statistical sense *(comments from a statistician needed here).*

   b) There is apparently a positive dose response across control, 800, 8,000 and 16,000 ppm groups, suggesting a different mechanism of carcinogenicity in

this range versus that of the 100 ppm group, in which the tumor incidence is elevated above that of the 800 ppm group. It is not unreasonable to consider that a different mechanism of carcinogenicity operates across such a vast dose range of 100 ppm to 16,000 ppm, which CARC has not acknowledged.

c) The presence of *carcinomas* which are absent in the control and in the 16,000 ppm group, despite the latter group's 96% adenoma incidence characterized in many cases by multiplicity, i.e. multiple adenomas.

d) Evidence in the low dose group versus the control of a more advanced stage of carcinogenesis in the *"natural history of neoplasia"* (foci of cellular alteration > adenoma > carcinoma) for hepatocellular tumorigenesis [44554901 (12/1/98) and Attachments 4 (5/4/98), 5 (5/29/98), 8 (4/1/99) (p. 2)] consisting of: I) presence of carcinomas; ii) incidence of carcinomas exceeds historical control in a weak historical control data base for 18-month studies (the large NTP data base consists of 2-year studies); iii) three (possibly four) incidences of multiplicity; iv) larger size of tumors; v) presence of liver masses as identified macroscopically versus none in control.

e) Even though, following the PWG which eliminated half of the carcinomas identified in the original study report (none of which were in the control), there remained eight mice with carcinomas in the dosed groups. Four of these were in the low dose group, and one (possibly two) of these mice in turn had two hepatocellular carcinomas. At one CARC meeting, our consulting pathologists said something to the effect: "There is something different about that low dose group, but I don't know whether its due to malathion." My response would be, this is a controlled study to evaluate effects of malathion, and findings in dose groups should be attributed to treatment, absent a definitive explanation.

Questions for the SAP Panelists

1) a) Should cholinesterase inhibition as observed in this study serve to conclude the 8000 and 16000 ppm dose levels were excessive? If the answer should prove to be yes, should the 800 ppm (ten-fold lower) dose level be considered adequate in spite of the fact it is well below the limit dose, however interpreted?

b) In the search for an MTD (or a dose level sufficiently high and challenging to provide an adequate carcinogenicity bioasssay), should cholinesterase inhibition, in and of itself, serve to set that mark, even in the absence of

cholinergic clinical signs; the concern being whether cholinesterase inhibition can serve as a roadblock for the testing of cholinesterase inhibitors at sufficiently high doses required for cancer assessment, as compared to other classes of compounds?

2) If it can be concluded dosing *was* excessive at 8000 and 16000 ppm, should the liver tumorigenic findings at these doses (both sexes) be discounted for use in quantitative risk assessment, absent any evidence the tumorigenic responses were due to anything other than the tumorigenicity of the test material [Attachment 18 (11/12/99]?

3) If findings at the top two doses are discounted because dosing was excessive, do the remaining dose groups, 100 and 800 ppm, constitute an acceptable study?

4) How should the findings in the low dose male group be interpreted?

5) Among males, do the characteristics of the liver tumor findings viewed in concert with the wide dose range (100 to 16000 ppm) suggest differing mechanisms of carcinogenicity across the dose range? Can panelists affirm PWG's conclusion that carcinomas in the lower dose groups can be discounted because they were not observed at the high dose level where adenomas were abundant?

6) Following the PWG, which downgraded half of the original 16 carcinomas in the male dose groups, can the historical control data base for the performing laboratory, which incorporates but 4 carcinomas total among males in the five studies, and not evaluated by this PWG, be considered meaningful in the interpretation? The NTP data base is for 2-year studies. Carcinomas in 18-month studies may be quite rare if diagnosed under a set of strict criteria.

II. Liver tumorigenic response in the female F344 rat

I am unable to accept the Pathology Working Group (PWG) report for the reasons indicated in my memorandum to the CARC Chairman [Attachment 29 (4/27/2000)]. These reasons are summarized and somewhat enhanced as follows:

1) Under PR Notice 94-5, a (PWG) convenes for the purpose of reading (diagnosing) histopathology slides, and not to perform an assessment of the carcinogenicity of the test material, as was done in this case. According to my understanding, the results of the histopathology re-reads should have been submitted for the *Agency's* use in assessment of carcinogenicity.

2) The PWG report says: "The purpose of the PWG review was to determine the

incidence of hepatocellular *neoplasms* (emphasis added) in female rats following currently accepted nomenclature and diagnostic criteria (Goodman DG, et al., 1994)." (p. 8). More correctly, in my view the PWG would be expected to determine incidences of lesions recognized as critical in the "*natural history of neoplasia*" (hepatocellular alterations, adenomas and carcinomas) as these would prove essential as "key events" and neoplasms to evaluate under EPA's draft 1999 Cancer Assessment Guidelines. An important observation from this study, before and after the PWG, in support of the concept of the natural history of neoplasia, and why all three of its elements need to be considered in the assessment under EPA's draft Guidelines is this. All carcinomas were downgraded to adenomas (and not another lesion), and when adenomas were downgraded, they were downgraded to hepatocellular alterations (and not another lesion). Hence, in this study, 5 carcinomas (the entire lot) were downgraded to adenomas, while 6 of the 8 original adenomas were downgraded to hepatocellular alterations, the "key events" in this drama.

In this case, it appears that final incidences of hepatocellular alterations (moderate degree of severity) are needed, as these lesions were selected out by the PWG Chairman to forward to the PWG as the most suspect among such lesions to be reclassified as adenomas by the PWG, and hence by that very selection qualify as "key events". Until the data are completely presented, an acceptable overall assessment of the neoplastic response under EPA's draft Guidelines is not possible. Consistent with this conclusion, an NTP pathologist recently advised me that when hepatocellular alterations appear in the absence of neoplasms in a study, they are not considered positive evidence of carcinogenicity; but, when hepatocellular neoplasms are present, the hepatocellular alterations are then employed in the assessment.

3) The study submission should provide in microscopic anatomic terms some explanation for the revisions of carcinoma to adenoma, and adenoma to hepatocellular alteration, to help the reader appreciate the level of difficulty or ambiguity in making these decisions. In other words, for example, what microscopic features led the study pathologist and the reviewing pathologist to concur on a diagnosis of a particular lesion as carcinoma, though later were not considered adequate to that diagnosis by the full PWG? This should be provided for each carcinoma and adenoma that was downgraded by the PWG. Also, this is particularly important for the 6000 ppm dose group, where 6 rats with hepatocellular alterations were identified (severity not recorded) by the PWG, as contrasted with 3 adenomas and one hepatocellular alteration (moderate) in the original assessment. *The concern is the extent to which the livers in the 6000 ppm group in question should now be considered to be no different from normal liver in the evaluation of the hepatocellular neoplastic response*.

4) Incidences of multiplicity (presence in one liver of two or more of the lesions constituting the natural history of neoplasia) should be clearly tabulated in the PWG report.

5) To the extent the PWG revised the original incidences of hepatocellular lesions (hepatocellular alterations, adenomas, carcinomas) as identified by the performing laboratory, one would expect the historical control incidences would also be vulnerable to revision under the scrutiny of this particular PWG, and thus until similarly evaluated the historical data base looses its relevance in the interpretation of this study. It may be that in the historical data base, all carcinomas would be downgraded to adenomas, and many adenomas downgraded to hepatocellular alterations under the criteria employed by this particular PWG. In other words, the possibility needs to be excluded that, in this peer review, the mark may have been set inordinately high for lesion classification, and might not be universally accepted among pathologists. In view of this concern, the PWG report should incorporate good reasoning why the historical data base should be considered relevant.

6) An additional interpretation of the slides (i.e. another opinion) should be obtained from pathologists from academia before the results of this PWG should be accepted.

7) The histopathology in the original study submission should be retained for risk assessment purposes until the above are satisfied.

Questions for the SAP panelists

1) Does the SAP accept the concept of a "*natural history of neoplasia*" for liver, involving hepatocellular alterations (or foci) > adenoma > carcinoma?

2) Should incidences of hepatocellular alterations (moderate degree of severity) be included (tabulated) in the PWG report for use as "key events" in evaluating the neoplastic response under EPA's draft Guidelines?

3) Should the PWG provide in anatomic terms information which explains the more critical differing conclusions among pathologists for the adenomas and carcinomas downgraded?

4) Should the more critical slides be examined by additional pathologists, e.g. from academia, to help answer the question of whether the criteria for tumor diagnosis employed by this PWG are accepted in the wider scientific community?

III. Liver tumorigenic response in male F344 rats

In my view, the study was inadequate for assessment of liver tumorigenic responses in the male rat at sufficiently high doses required for an acceptable negative study in a cancer bioassay. The concern actually extends to all potential neoplastic responses, i.e. not confined to the liver.

It may be the F344 rat is an inappropriate model for assessing carcinogenicity of malathion in the male gender. Along these lines, I should note that CARC concluded 500 ppm to be an acceptable dose level, and thereby the study to be an acceptable negative study in males. This is to be contrasted with the acceptable dose level of 6000 ppm (over ten fold-higher) in females. The earlier 1978 NCI study in Osborne-Mendel rats employed doses of 0, 4700 and 8150 ppm (twa), which were not found to be excessive either in terms of survival, or for any other reason, in either sex. I am concerned that in discounting the top two doses (6000 and 12000 ppm), the "power" of the current study has been unacceptably compromised in terms of the dose being too low for proper assessment of malathion, and having in effect been reduced to a two dose study, in which all tissues other than liver, lung, kidney, target organs were not examined histopathologically in the 500 ppm group, as would be required should it serve as the high dose group for the study. The lowest dose level, 100/50 ppm, was chosen as I understand in search of a NOEL for cholinesterase inhibition. In other words, 100/50 ppm is an extremely low dose for malathion, bordering on the LOEL/NOEL for cholinesterase inhibition, and 500 ppm is not that much higher. Under 40 CFR the majority of food tolerances for malathion are 8 ppm.

Question for the SAP panelists

Should the dose level of 500 ppm be considered adequate for carcinogenicity testing in male rats, and hence the study accepted as satisfying requirements for a negative study for liver carcinogenicity (or any other neoplastic response) in that gender?

IV. Thyroid C-cell tumorigenic response in male F344 rats

In my view, the weight of the evidence supports as positive thyroid C-cell carcinogenicity.

The 1990 Office of Pesticide Programs' Health Effects Division Peer Review of the carcinogenicity of malathion identified as principle tumorigenic findings of concern in the existing data base those of the thyroid (C-cell ***and*** follicular cell), adrenal gland (pheochromo-cytoma) and mononuclear cell leukemia in the rat, and hepatocellular tumors in the mouse. The committee elected to require additional carcinogenicity bioassays of malathion and malaoxon, Dearfield (1990) (4/12/1990). In order to understand more fully the character of the data base existing at the time of the 1990 peer review, reference is made to other documents cited in Dearfield (1990), including those of Dementi (1990) (4/20/1990), the 1978-79 National Cancer Institute (NCI) bioassay reports, and Huff et al (1985) (a Pathology Working Group assessment of the earlier NCI studies, which I might add was largely focused on the thyroid response). *Suffice it to say,*

thyroid neoplasia was a principle matter of concern at the 1990 peer review, and this earlier data base should be factored into the current assessment of the thyroid carcinogenic response. Furthermore, given the earlier data and concern, there is in my view a peculiar incumbency that the newer studies address the thyroid tumorigenic concern in a definitive manner. However, for reasons presented here, thyroid tumors (C-cell and follicular cell) remain a positive finding in the more recent study of malathion in the rat. *It is imperative that individuals evaluating the data base read the references pertaining to the earlier NCI studies, particularly those of Dearfield (1990), Huff et al (1985) and the HED review (Dementi, 1990) where the assessment of the thyroid neoplastic findings are concerned.*

In the more recent (1996) study of malathion in the F344 male rat, the incidence of C-cell carcinoma was clearly a numerically and statistically significant finding in the 0-500 ppm dose range, as disclosed in Tables 10a and 10b of the CARC report (pp. 21-22). This finding was observed in the dose range considered acceptable by the CARC. As indicated previously, the thyroid C-cell response was of specific concern in the earlier studies, as noted in Dearfield (1990). Perhaps most noteworthy was the positive C-cell response in the 1978 NCI malaoxon study in the F344 rat, which was statistically significant by dose trend and high dose pairwise comparisons in *both sexes*, as identified by the PWG [Huff et al (1985); table 5, p. 168]. In my view that study was clearly positive for this neoplastic response. However, the finding was not confirmed in the more recent malaoxon study in the F344 rat, though I personally question the acceptability of that study at the top two dose levels, as discussed in this assessment under item X) (Leukemia in the malaoxon F344 rat study). It is often mysterious why a particular finding is not duplicated in a repeat study. This was seen in the case of C-cell tumors with malaoxon, and it was also seen in the B6C3F1 female mouse liver tumor response, which was absent in the original NCI study, but remarkable in the recent bioassay. It is my understanding, positive studies take precedent, absent a good reason to the contrary.

In the recent malathion study, as observed in one of the supporting documents [Attachment 10 (5/18/99)], the response across the 0-500 ppm dose range not only yielded increased carcinoma, but emulated a type of response described by McConnell, et al (1986), where the increased carcinoma incidence presented possibly as a progression from adenoma, such that the combined adenoma plus carcinoma incidence in the dose group was not increased. McConnell et al (1986) say for example: *"A particular chemical that induces a statistically significant shift in tumor expression from benign to malignant without the total incidence increasing may be regarded as a carcinogen. However, when the benign and malignant neoplasms are combined in such a case, the study would be classified as negative, implying that the chemical is non-carcinogenic. Thus combining neoplasms in this case would result in a false-negative effect." (p. 283)* Therefore, I am concerned over CARC's decision to rely upon the combined (adenoma plus carcinoma) tumor response, which ignores the evidence of progression to malignancy, to discount findings in the 500 ppm dose group. There was also multiplicity in the 500 ppm group, manifested as *two* instances where both a C-cell adenoma and carcinoma were present. The incidence of C-cell tumors declined beyond this dose level in a *dose related manner* at 6000 and 12000 ppm dose levels, dose levels considered excessive by the

CARC due to high mortality (74% at 6000 ppm and 100% at 12000 ppm). My contention here is that the expression of a positive tumorigenic response across the 0-500 ppm dose range, was not manifested further at the more than *ten-fold higher* dose levels of 6000 ppm and 12000 ppm dose levels due to both increased mortality (which substantially exceeded 50% in both groups) and competing toxicity. The data suggest that a more dramatic tumorigenic expression of the sort observed at 500 ppm might be anticipated to occur, indeed peak out, somewhere just on either side of 500 ppm, i.e. in a range not adequately tested. *Negative findings at excessive doses cannot be accepted to characterize the test material as negative for carcinogenicity, particularly when evidence of a positive response (trend, pairwise comparison, multiplicity, malignant tumors, substantial exceedance of historical incidence) exists in the acceptable lower dose range.* For the sake of argument, if one knew only of the findings at 0, 6000 and 12000 ppm, one might conclude the study to be negative for thyroid C-cell tumorigenicity. But given that both doses were excessive, the study would not qualify as an acceptable negative study. Now to the extent the study would not be considered acceptable as a negative study at 6000 and 12000 ppm, I would find it egregious, particularly from the perspective of public health protection, to employ findings from such a study to actually discount the positive findings observed in the acceptable dose range, 0-500 ppm. Indeed, positive findings in a dose range much more relevant and of greater concern in terms of assessing human risk is being compromised in the interpretation of carcinogenicity.

To conclude as the CARC has done that sufficient animals were at risk beyond week 43 in the higher dose group, where the first tumor was observed, places too much reliance upon what could have been spurious timing of but one such response, while ignoring the impact of competing toxicity and increased mortality on those animals that may have survived beyond week 43. It also assigns preferential weight to the findings at dose levels which the CARC itself considers excessive. Clearly, the C-cell carcinoma response in this study, taken in conjunction with findings in the earlier NCI studies, constitutes positive evidence of carcinogenicity, indeed at low doses. The idea behind high dose, MTD but not above, testing is to enhance the power of animal bioassays to detect possible carcinogenicity that otherwise may not be detectable at low doses.

Conservatively, and in the interest of protecting the public health, to the extent the positive tumorigenic (carcinoma) response at 500 ppm is discounted, the study should be considered unacceptable to evaluate C-cell carcinogenicity in the male F344 rat, due to excessive mortality and competing toxicity at 6000 and 12000 ppm. [Attachments 17 (10/28/99), 18 (11/12/99), 21 (2/7/00), 24 (4/24/00)].

Questions for the SAP Panelists

1) Is it appropriate to conclude the top two doses excessive, and then, in effect, treat the negative findings at such doses as acceptable in discounting positive findings at lower doses (indeed more than ten fold lower) considered acceptable?
2) Is the C-cell carcinoma response to be concluded real and an appropriate end point on which to base a conclusion regarding carcinogenicity?

V. Thyroid follicular cell tumorigenic response in male F344 rats

According to my interpretation, evidence of a thyroid follicular cell neoplastic response cannot be discounted if the study is to be considered acceptable.

As indicated in my previous discussion of thyroid C-cell tumors, follicular cell tumors were at issue in the earlier NCI studies of malathion in the Osborne-Mendel rat and malaoxon in the F344 rat, again indicating the thyroid to be a vulnerable tissue. The reader is directed to the HED review (Dementi 1990) for discussion of the follicular cell neoplastic findings in the earlier NCI studies.

In the current study as indicated in Table 9 (p. 20) of the CARC report, for male rats there was a positive dose trend ($p = 0.035$) and a near positive pairwise comparison for the 6000 ppm group versus the control ($p = 0.077$). It is noteworthy that there also were two carcinomas each in the 500 ppm and 6000 ppm groups, as contrasted with none in the control. Again, in view of the fact that mortality was 100% in the 12000 ppm group, and taking into consideration the influence of competing toxicity, it is not particularly surprising the incidence of this tumor did not increase further in this high dose group, and the fact that it didn't at the clearly excessive dose of 12000 ppm cannot be used to discount the finding across the lower dose range. Also, arguably, the increase observed at 6000 ppm, though nearly achieving statistical significance ($p = 0.077$), may have been under-expressed for the same reasons, increased mortality (74%) and competing toxicity. ***So, in my view, in light of the principles of interpretation set forth earlier in this document, the emerging neoplastic response, already evident in the data (positive trend, near positive pairwise comparison), cannot be discounted if the study is to be considered acceptable in male rats.*** [Attachments 18 (11/12/99, 21 (2/7/00)]. Furthermore, in consideration of this particular tumorigenic finding as observed in the data base as a whole, rationale exists to conclude as positive a follicular cell neoplastic response.

Questions for the SAP Panelists

1) In consideration of the evidence for a follicular cell neoplastic response evident in the earlier NCI studies, taken in concert with the response in this study, can this study be concluded to be negative and to have satisfied the earlier concern?

2) To the extent the study is considered acceptable in males, should the follicular cell findings be concluded as positive evidence of a neoplastic response?

VI. Nasal tissue tumorigenic response in the F344 rat

As recorded in the CARC report, following an external peer review of nasal tissue histopathology there were four extremely rare nasal adenomas, one in each sex at both the 6000 ppm and 12000 ppm dose levels. According to my understanding, extremely rare or unusual neoplasms do not require statistical significance or a clear dose response to be interpreted as constituting positive evidence of carcinogenicity of a test material. *In my view, CARC has marginalized this finding in the overall carcinogenicity assessment.[Attachments 13 (7/123/99), 19 (12/7/99), 20 (1/12/00),* 24 (4/24/00)]

I accept these tumors as evidence of carcinogenicity, based primarily upon their extreme rarity and absence in controls. As I understand, CARC discounts both findings in males, and the one at 12000 ppm in females, predicated upon excessive dosing. I have the same concern as expressed in reference to other neoplastic findings seen at doses considered excessive, namely, "...positive results obtained above the EMTD are acceptable as evidence of carcinogenicity unless there is convincing evidence to the contrary." However, as with certain other neoplastic findings in this study, to the extent these are discounted in the high dose groups, the study is unacceptable in males. The issue is complicated by evidence of nasal histopathology in the long term combined chronic toxicity/carcinogenicity studies in the F344 rat for both malathion and malaoxon, and in the dose range-finding and subchronic inhalation studies of malathion in the rat. In the malathion study, CARC concluded that in both sexes, lesions of the olfactory/ respiratory mucosa were more severe at 500, 6000 and 12000 ppm. (p. 30). Noteworthy is the fact that like liver, the nasal mucosa has remarkable metabolic capability toward xenobiotics. To the extent the nasal tumors are secondary to the toxicologic effects on the nasal mucosa, such histopathology may be a "key event" under EPA's draft cancer Guidelines. While a new subchronic inhalation study is being required, I am not satisfied that CARC has an adequate interim handle on risks, including carcinogenicity, posed with respect to the nasal mucosa, particularly by the inhalational route of exposure. *Mentioned at one time was the possible need for a carcinogenicity study by the inhalational route of exposure*.

Nasal tissue vulnerability is an important and unresolved issue at this time. Other noteworthy issues include:

1) The somewhat confusing histopathology diagnosis of the nasal tumor of a male rat of the high dose group, which as explained in supporting documents, has been variously characterized in nasal turbinate section 2 as "carcinoma", "hyperplasia" and finally "adenoma", arising in section 1 and extending into sections 2 and 3.

2) The CARC report (p. 18) describes the nasal tumor of the olfactory epithelium in the 6000 ppm dose group male as **"***esthesioneural epithelial neoplasm***"**, a term new to the CARC at its last meeting. Both Drs. Henry Bolte and James Swenberg have identified the lesion as "nasal mucosa (olfactory): adenoma". As toxicologist, I am not aware of any submission from the registrant in which the diagnosis for

this tumor was revised or clarified. The CARC seems to be employing its own characterization of this tumor such that it becomes treated as not combinable with the other three nasal adenomas. I must question CARC's terminology for this tumor, as the tumor has not been so characterized by any examining pathologist. Furthermore, nasal histopathology was so altered in this study that olfactory neural tissue was "replaced by ciliated and nonciliated columnar epithelial cells", which in turn, could have been the actual tissue from which the adenoma arose, even though located in the olfactory region of the nasal passages *(comment from expert pathologists is needed here*). In other words, the pathology sheets do not state that the lesion was of neural tissue origin, or that it was a neural tissue lesion, only that it was located in the olfactory region of the nasal passages.

3) As to the question of whether all four tumors should be combined, my view is they should be combined. However, if indeed one should prove to be of a different classification, it nonetheless represented a response in common with the other adenomas to a remarkable insult to the nasal mucosa, and should not be interpreted as posing any less concern than would four identical tumors.

It is noteworthy that in the 1978 NCI Osborne-Mendel rat carcinogenicity study (NCI Carcinogenesis Technical Report Series No. 24) (malathion doses: 0, 4700 and 8150 ppm), in which study nasal tissues were *not examined histopathologically*, there was a nasal cavity tumor identified grossly in a female rat of the high dose group, which was subsequently confirmed by microscopic examination as carcinoma. I confirmed by an NTP spokesman that this was the only nasal section examined microscopically, and that nasal slides were not taken, for this or the other malathion studies.

I remain particularly concerned over the finding of a total of ***eight exceedingly rare tumors*** [4 nasal adenomas, 4 oral squamous cell tumors (discussed below)] found only in dosed groups as identified in the nasal histopathology evaluation, an evaluation which did not completely evaluate the oral tissues. Also, none of these eight tumors was seen grossly as was the nasal carcinoma from the older malathion study mentioned above. It is unfortunate nasal and oral cavity tissues were not examined histopathologically in the earlier NCI studies

Questions for the SAP Panelists

1) Should any of the four nasal tumors be discounted for reasons of excessive dosing?

2) Should the term "esthesioneural epithelial neoplasm" be accepted as correct nomenclature for the nasal lesion diagnosed as "nasal mucosa (olfactory): adenoma" by the study and reviewing pathologists? Does this make any substantial difference in the assessment of carcinogenicity of the nasal mucosa in

this study?

3) Given that nasal pathology was more severe across the top three doses, does this observation pose a concern for tumorigenicity at 500 ppm, even though tumors were not identified in that dose group?

4) Given the extreme rarity of the nasal tumors in F344 rats, do these tumors constitute adequate evidence of carcinogenicity? If not, what level of concern should be ascribed to these lesions?

5) If as indicated in the CARC report the nasal tumors in females occurred in section 5, where little other histopathology was noted, how might this influence the level of concern?

6) Is carcinogenicity testing by the inhalational route indicated?

VII. Oral tissue squamous cell tumorigenic response in the F344 rat

In my view, absent any further histopathology assessment, the oral tissue squamous cell tumors constitute positive evidence of carcinogenicity that cannot be discounted.

Questioned here are both the adequacy of the oral cavity histopathology and CARC's conclusion regarding the tumorigenic response [Attachments 13 (7/13/99), 14 (7/22/99), 15 (9/21/99), 19 (12/7/99), 20 (1/12/00)]. These particular oral squamous cell tumors were identified *incidentally* in the histopathology assessment (and re-assessment) of *nasal* turbinate tissues. In my view, once these tumors were identified among the limited oral tissues (palate, alveoli of teeth) examined in a nasal histopathology assessment, CARC and/or the registrant should have pursued histopathology on all oral tissues, which were not in fact examined in this study. The CARC elected not to require oral cavity histopathology.

I consider more compelling an argument to require oral histopathology than the nasal tissue re-examination that was required by CARC, since the oral tissues, unlike the nasal tissues, were not examined in the original study, and in the face of three exceedingly rare squamous cell tumors (two malignant ones) among females, in which gender there is less concern over excessive dosing. Also, two of the lesions (one of each sex) are seen in the lowest dose group, a dose level of particular concern with respect to the public health.

In various memoranda and in presentations to CARC, I have labored to explain that squamous cell tumors of the *palate* are so rare in NTP's data base, that I have been unable to confirm even a single incidence in male or female F344 rats. [Attachment 14 (7/22/99)] Although very rare, squamous cell tumors are reported in NTP's data base for other regions of the oral cavity, most notably the tongue and oropharynx, but these tissues were not examined in the

malathion study.

The oral cavity squamous cell tumors (three in females, one in males) cannot be discounted as evidence of carcinogenicity. The squamous cell carcinoma of the squamous epithelium lining the alveolus of a tooth in the female low dose group should be treated as combinable with the squamous cell tumors of the palate, as the cites are contiguous and of the same epithelium, so much so that the squamous cell carcinoma of the 12000 ppm group was originally diagnosed as an alveolus squamous cell carcinoma, but revised by peer review as being more appropriately a squamous cell carcinoma of the palate. Furthermore, as these tumors were identified in but a partial and inadequate assessment of the oral cavity histopathology, they should take on even more weight in CARC's assessment of carcinogenicity.

As indicated previously for the nasal tumors, rare tumors, particularly those which are exceedingly rare, need not be statistically significant nor characterized by a positive dose response to be considered positive evidence of carcinogenicity, as I understand. However, add to that the fact that in females there was evidence of progression in the form of a squamous cell papilloma of the palate at 6000 ppm, and a squamous cell carcinoma of the palate at the 12000 ppm dose level, constitutes an added matter for consideration before the CARC, which has been ignored or marginalized. Furthermore, among females, the carcinoma at the high dose level should not be disregarded, as there is no evidence it was due to excessive toxicity or anything other than the tumorigenicity of the test material (EPA's Guidelines, as cited in Attachment 18)

I find the following statement in the CARC report to be unacceptable, to say the least: *"However, the CARC* ***concluded*** *(emphasis added) that additional pathological evaluation would not alter their conclusion." (p. 21),* unless it was rendered at the time CARC had already concluded malathion a "likely human carcinogen". If that were the reason, it would no longer apply.

Questions for SAP Panelists

1) Given that the oral squamous cell tumors were identified incidentally as seen in the palate and alveolus of a tooth in the conduct of a nasal tissue histopathology evaluation, what level of confidence is there that even these tissues received adequate histopathology evaluation, i.e. at least equivalent to that which would be expected from an actual oral cavity assessment?

2) Should OPP require oral cavity histopathology?

3) Given the rarity of the oral tumors (particularly of the palate) in F344 rats, do these tumors constitute adequate evidence of carcinogenicity? If not, what level of concern should be ascribed to these lesions?

4) Though the nasal and oral tumors are of different cell type, given their locations

and in their rarity do they denote a common concern for carcinogenicity?

5) Is carcinogenicity testing by the inhalational route indicated to evaluate oral cavity responses? In relation to this question, Dr. Henry Bolte, study pathologist, has said in correspondence that it is in inhalation carcinogenicity studies that both nasal and oral cavities receive histopathology assessments.

VIII. Interstitial cell testicular tumors in the F344 rat

In my view, this neoplastic response was positive as evaluated by the appropriate statistical method of analysis and cannot be discounted either on the grounds that the finding is an artifact of the statistical procedure employed, or that this neoplastic response is of questionable importance.

This tumorigenic response was statistically significant across all four doses as presented in the study report, and was positive across the top three doses as analyzed by the Peto test within HED. I should note this was the prescribed test. These statistical analyses reveal a dosing related higher incidence than expected of this tumorigenic response, indicating a positive carcinogenic effect by a recognized definition of a carcinogen.

As to the interpretation, the literature indicates that incidences as high as 100% for this tumor type occur *near term* in normal control F344 rats. I am not aware of any evidence that they appear *ubiquitously* prior to near term. CARC's argument is that because the treated animals died early, there may have been early observation of the tumor. This is probably true, though the interpretation may not be correct, i.e. that the tumors were seen early simply because they were examined at an earlier time point, where they normally occur. On the other hand, perhaps they were seen, *ubiquitously*, at the earlier time points because there was a compound related earlier onset. The Peto test must be viewed as the appropriate test to establish the latter alternative as real, at least until *proved* otherwise, where the burden of proof resides with the registrant. The CARC has indicated the finding as an artifact of the Peto Test, and that "There was no serial sacrifice to determine latency." (p. 27) However, mechanistic data of this sort is the duty of the registrant to provide if he wishes to have the positive findings discounted. In my view, it is inappropriate for CARC to note the absence of the needed data, and then conclude the findings are not real in the face of the needed data. CARC also says: "this tumor type is not useful in overall evaluation since its occurrence is similar at all dose levels" (p. 27), a kind of circular reasoning.

It has been a curiosity just why the Peto Test should have been performed, if the findings, clear as they are, can be so arbitrarily discounted and disregarded. The fact remains it is positive by all methods of statistical analysis at the top three doses, and by certain methods extends to the lowest dose, as very persuasively presented by the study's own statistician, indicating a biological effect of the test material even at doses as low as 100/50 ppm. Furthermore, regardless of the

inherent importance to be assigned to this particular tumor type, a positive neoplastic response in any tissue constitutes evidence of carcinogenicity, and speaks to the wider concern of the carcinogenicity of the test material.

A published work has been cited (McConnell et al 1986) indicating this tumor type can progress from the benign to the malignant stage. Therefore, viewed in terms of human experience, earlier onset would pose enhance opportunity for this progression to occur. This prospect would run counter to CARC's claim that this is a non-lethal tumor, and hence to the implication of a lack of concern for this tumor type. [Attachments 11 (6/7/99), 24 (4/24/00) (p. 6)]

Questions for the SAP Panelists

1) Is the Peto Test the correct test for statistical analysis of this data, and if so, should the statistically significant findings across at least the top three doses, and possibly the lowest dose, be considered treatment related? If so, does it qualify as adequate evidence of carcinogenicity based on decreased latency or enhanced development of the response?

2) Should the answer to the above questions be yes, is this finding to be considered as of little importance in terms of human risk considerations, i.e. not a very important tumor type, as suggested by CARC? Stated differently, how should the finding be characterized in the overall assessment of carcinogenicity?

3) If this finding should be considered as of low rank in the evaluation of carcinogenicity potential, does it not identify a biological effect of the test material certainly at the top three doses, and possibly so at the lowest dose level?

IX. Leukemia response among male F344 rats under OSTP (1985) definition of carcinogen

Claimed here is that a dose related increased incidence of mortality *attributed to leukemia* among male rats *diagnosed with leukemia* constitutes positive evidence of carcinogenicity under the second aspect of OSTP's definition of carcinogen, namely, "..... or significantly decreases the time it takes a naturally occurring (spontaneous) tumor to develop relative to an appropriate background or control group. Either phenomenon is said to represent the effects of a carcinogen." (pp. 10410-10415)

Leukemia has been a recurrent tumorigenic end point in the malathion data base. It was one of the neoplasms considered by a PWG (Huff et al 1985) conducted on the earlier NCI study in the F344 rat, and is discussed in Dearfield (1990).

I contend the dose-related increased mortality (where mortality itself indicates a more

advanced stage) is evidence of a dose-related increased rate of development of leukemia. [Attachment 9 (4/27/99)] In this particular study, the number of male rats among 55 rats per group diagnosed with leukemia (death due to leukemia) were 23(7), 16(7), 24(14), 18(13) and 1(1) for the control, 100/50, 500, 6000 and 12000 ppm groups, respectively. Hence, among rats diagnosed with leukemia, the percentages dying with leukemia were: 7/23(30%), 7/16(44%), 14/24(58%), 13/18(72%) and 1/1(100%). CARC argued at one meeting that rats harboring leukemia are simply more susceptible to early death due to the increasing secondary toxicologic burden of the test material, but to confirm that possibility and to discount the possibility of a direct compound effect in development of the response, the mechanism would need to be established, which is a responsibility of the registrant. I should note the effect is manifested even at the lower doses, i.e. dose levels CARC considers acceptable. CARC has also sought to allay this finding on the grounds that rats dying with leukemia died no sooner in dosed groups and that a diagnosis of death due to leukemia is subjective. (p. 29 of CARC report). I have responded that a) leukemia in the F344 rat is of late onset, and the time interval between onset of the disease and death is so truncated that it is difficult to establish earlier time to death, but nonetheless the data I presented do suggest that rats in the 6000 ppm group died earlier; b) diagnosing death due to leukemia is not that subjective, and is an important responsibility of the study pathologist. [Dementi to Burnam (6/8/2000)]

Questions for the SAP Panelists

1) Do these findings constitute meaningful evidence of a dosing related increased development of leukemia among animals that have leukemia? If so, does this constitute positive evidence of carcinogenicity under the second aspect of the OSTP (1985) definition of carcinogen, namely that of enhanced development or progression of the disease?

2) If not, do these finding constitute evidence of a compound related biological effect extending even to the lowest dose level?

X. Leukemia in the malaoxon F344 rat study

The final draft of the CARC report acknowledges the positive leukemia response among male rats, trend ($p = 0.03$), and pairwise comparison for the high dose group ($p = 0.05$). I accept this as positive evidence of carcinogenicity. The pairwise comparison for the intermediate dose group was marginal ($p = 0.07$). The CARC concluded leukemia as not treatment related ".....since statistical significance was seen only in males at a dose that was determined to be excessive, there was no dose response, and the incidences were within the historical control range....." (p. 35 of the CARC report). The CARC report goes on to say leukemia was not seen in the earlier NCI studies. [Attachment 24 (4/24/00) (p. 2)]

The problems I have with this conclusion include the following. As with certain other

tumorigenic findings in the malathion data base, I do not accept discounting the high dose positive finding as treatment related, *unless it can be shown to be a result of toxicity as opposed to the tumorigenicity of the test material*. It is more defensible to accept as positive studies exhibiting effects at excessive doses than to accept as negative a study absent effects at excessive doses. Furthermore, in the face of a positive trend and the increase in incidence at the high dose, I question discounting the increased leukemia in the 1000 ppm dose group, where there was marginal statistical significance ($p = 0.07$). Mortality among male dose groups in this study were: control (29%), 1000 ppm (42%, N.S.) and 2000 ppm (53%, statistically significant). Since leukemia in F344 rats is of late onset [Dementi to Burnam (6/8/00)], increased mortality may have compromised expression of leukemia in a study where incidence in both these study groups was already marginally significant. That leukemia expression can be compromised at high doses due to increased mortality and/or competing toxicity is well illustrated in the malathion study, where incidences among 55 male rats per group were: 23, 16, 24, 18 and 1 for the control, 100/50, 500, 6000 and 12000 ppm dose groups, respectively. The fact that leukemia expression was essentially ablated at the 12000 ppm dose level supports the argument of not accepting as negative a study absent effects at excessive doses. Likewise in females, mortality was: control (13%), 1000 ppm (44%, statistically significant) and 2000 ppm (49%, statistically significant), such that leukemogenic potential may have been compromised *in some degree* by enhanced mortality.

Where assessment of leukemogenic potential in females is concerned, both of the high dose group levels may have been excessive, and hence the study may be inadequate to serve as a negative study for leukemia in females. CARC has concluded the 2000 ppm dose level was excessive, while I have argued the 1000 ppm group was also excessive in females, by CARC's criteria, namely: a) increased mortality [mortality was 44%, very nearly the same as that of 49% at 2000 ppm (though not mentioned in the CARC report, see p. 35)] and b) severe cholinesterase inhibition (p. 35 of CARC report).

The CARC report indicates there were no findings in the earlier NCI studies, however according to Huff et al (1985) concerning a PWG performed on the earlier studies, there was similarly among male F344 rats in a malathion study, a significant increase in leukemia in the low dose (2000 ppm) group but not in the high dose (4000 ppm) group. Huff et al (1985) say reduced survival [control (54%), 2000 ppm (28%), 4000 ppm (0%)] made interpretation of leukemia incidence difficult, saying life-table analyses suggest increases in leukemia, primarily in the low dose group (statistically significant), but life table analysis was discounted because leukemia was not a cause of death (death attributed to "chemical toxicity"), while the increase was not significant by incidental tumor tests or Fishers Exact Test, and, hence, not related to treatment. However, in the recent malathion F344 rat study, leukemia was a principal cause of death, and a similar interpretive problem exists as to whether the study satisfies as a negative malathion study for leukemia among male F344 rats.

Contrary to what CARC claims, a positive dose trend is evidence of a dose-related effect.

The historical control range given is for animals in control groups of normal survival, where

leukemia in the F344 rat is late occurring, and may not be applicable to compare incidences of leukemia diagnosed in rats that died prematurely.

Questions for the SAP Panelists

1) Should the positive (statistically significant) findings for leukemia among male rats be concluded not treatment related?

2) What can be said about the leukemogenic potential for malathion and malaoxon as gleaned from the various studies under review?

XI. Mutagenicity of malathion

I have concerns over changes in the mutagenicity assessment for malathion as reflected during the recent course of time in the draft and final CARC reports.

The September 20, 1999 draft CARC report says in a concluding statement: "The positive mutagenicity studies with malathion support the evidence of liver tumor induction in male and female mice. Based on the overall results, there is a clear concern for somatic cell mutagenicity. No further testing is required since all mutagenicity issues have been addressed." (p. 21)

The February 2, 2000 draft CARC report says: "The Committee concluded therefore, that while the evidence for mutagenicity as an influence on the carcinogenicity of malathion is weak, at this time it can not be ruled out." (p. 41)

The final April 28 report says: "Although the structure of malathion suggests electrophilicity, the Committee concluded that the weight of the evidence supports neither a mutagenic hazard nor a role for mutagenicity in the carcinogenicity associated with malathion." (p. 37)

To the extent that mutagenicity of a test material enters into the weight-of-the-evidence for the carcinogenicity classification, it is essential the mutagenicity conclusion reflect the findings in the data base. It is my understanding there has been little if any change in the mutagenicity data base between September 20, 1999 and April 28, 2000. Therefore, the substantial revision in the conclusion that has taken place during this time appears to reflect differing interpretations of the existing data base.

Question for the SAP Panelists

Does the Panel confirm the final CARC report conclusions concerning the mutagenicity of malathion?