CHPA CONSUMER HEALTHCARE PRODUCTS ASSOCIATION
founded 1881

January 27, 2020

Mr. Julian Leichty
Office of Environmental Health Hazard Assessment
Proposition 65 Implementation Program
P.O. Box 4010, MS-12B
Sacramento, California 95812-4010

Via electronic submission

Re: Request for Public Comment on Hazard Identification Materials for Acetaminophen

Supplemental Comments of the Consumer Healthcare Products Association on the Statistical Analyses of the Tumor Data from the Amo and Matsuyama (1985) Study of Acetaminophen in Mice

Dear Mr. Leichty,

This submission supplements the previous comments of the Consumer Healthcare Products Association (CHPA) dated November 4, 2019. Our initial comments noted that Amo and Matsuyama (1985) reported no statistically significant increase in any tumor in male or female mice administered up to 6000 ppm of acetaminophen in the diet for up to 134 weeks. The study authors concluded: "The results of the present tests show that feeding the maximum tolerated dose of acetaminophen (0.6% diet) held no carcinogenic hazard for B6C3F1 mice."[1] IARC (1990) reached a similar conclusion about this study: "No difference was found in the incidence of tumours at any site between treated and control mice (Amo and Matsuyama, 1985)."[2] In contrast, the Hazard Identification Document (HID)[3] re-analyzed the Amo and Maytsuyama (1985) study and reported statistically significant increases in hepatocellular adenoma and carcinoma combined and in pituitary gland adenoma among high-dose female mice.

At the time of our submission in November, we were not aware of the reasons for the conflicting statistical results; however, we noted that the study authors' conclusion that acetaminophen poses no carcinogenic hazard for B6C3F1 is consistent with the results of other carcinogenicity studies of acetaminophen in mice, including the NTP cancer bioassay. We have since evaluated the statistical methods used to evaluate the tumor data, and the purpose of this supplemental submission is to explain why the statistical results described in the HID differ from those of the study authors.

[1] Amo, H., Matsuyama, M., 1985. Subchronic and chronic effects of feeding of large amounts of acetaminophen in B6C3F1 mice. Nihon Eiseigaku Zasshi 40, 567-574.
[2] IARC, 1990. Monograph for Paracetamol (Acetaminophen). Volume 50. World Health Organization, International Agency for Research on Cancer, Lyon, France, pp. 307-332.
[3] HID (2019) Evidence on the Carcinogenicity of Acetaminophen, September.

The HID does not mention the results of the Amo and Matsuyama (1985) study as provided by the study authors.

The HID on page 5 states: "In the female B6C3F1 mice exposed to acetaminophen in feed for up to 134 weeks (Amo and Matsuyama 1985), the incidence of hepatocellular adenoma or carcinoma combined was significantly increased in the high-dose group by pairwise comparison with controls, with a significant dose-related trend."[4] However, the HID does not mention that the study authors found no statistically significant increase in any tumor type. It is unclear that the statistical results presented in the HID are not those of the study authors, but rather a re-analysis of the data by OEHHA using different statistical methods. The only indication that the HID is re-interpreting the results of the study appears in a footnote in Table 15 on page 116, which states:

> "Footnotes for the studies by Amo and Matsuyama (1985):
> The denominator is the number of animals alive at the occurrence of the first tumor at any site, which was 51 weeks.
> *p < 0.05, pairwise comparison with control by Fisher's exact test (performed by OEHHA) Trend p-value: exact trend test (performed by OEHHA)"

If a reader did not notice this footnote in Table 15, which appears 111 pages after the study results are first described in the HID, there would be no way to know that the findings of the Amo and Matsuyama (1985) study presented in the HID are not those of the study authors themselves, but instead a re-analysis by OEHHA.

Why do the statistical results presented in the HID differ from those of the study authors?

The HID used the Fisher exact test to conduct a pairwise comparison between the high dose females and the controls. For this test, there are several important decisions that must be made. First, a decision must be made about which animals to include in the analysis. It is common to use the number of animals at the outset of the study as the denominator. For example, NTP typically chooses the number of animals assigned to the study for its statistical evaluation of tumor data. In comparison, OEHHA chose "the number of animals alive at the occurrence of the first tumor at any site, which was 51 weeks."[5] This approach is not standard practice.

Second, a choice must be made whether to use a one-tailed or a two-tailed test. A one-tailed test is used when the investigator cares only about increases in tumors. A two-tailed test is more appropriate when both increases and decreases in tumors are of interest. Although it is not stated in the HID, OEHHA chose to use a one-tailed test. This approach does not account for the reality that some test substances can decrease the incidence of tumors, as has been reported in NTP bioassays. In fact, Amo and Matsuyama (1985) noted slightly lower rates of onset of tumor-

[4] Id., p. 5.

[5] Table 1 of Amo and Matsuyama (1985) describes both the number of mice at the start of the study and the "effective number of mice." A footnote in Table 1 for the "effective number of mice" states: "Survived more than 51 weeks, when the 1st neoplasm was found." Figure 7B in Amo and Matsuyama appears to show that the first tumor among the females was seen around week 51 in a low dose female. In comparison, Figure 7A shows that among males, the first tumor appeared around week 48 in a control male. Apparently, the data in Table 1 describes the effective number of mice that survived more than 51 weeks for each group of both male and female mice, and these were the values used by OEHHA for its statistical re-analysis.

1625 Eye Street, NW, Suite 600 | Washington, DC 20006 | T 202.429.9260 | F 202.223.6835 | chpa.org | @chpa

bearing (tumors of all sites combined) among female mice at both dose levels of acetaminophen compared to controls.[6]

Third, a decision must be made about the appropriate p-value to achieve statistical significance for tumor data in an animal study. The NTP generally regards as carcinogenic any chemical that produces a high-dose increase in a common tumor that is statistically significant at the 0.01 level or a high-dose increase in an uncommon tumor that is statistically significant at the 0.05 level.[7] Liver tumors are common tumors in male and female B6C3F1 mice, the strain used by Amo and Matsuyama (1985). In fact, NTP has reported that liver tumors are the most common spontaneous tumor observed in its cancer bioassays using this strain of mice, and it is also the most common tumor type reported to be significantly increased by substances tested in mice in NTP bioassays. Thus, by NTP's criteria, it is appropriate to use a p-value of 0.01 to determine whether an increase in a common tumor (such as mouse liver tumors) is statistically significant. Similarly, the FDA uses different p-values for statistical significance for common and uncommon tumors in animal carcinogenicity studies.[8] OEHHA chose to use a p-value of 0.05, which is not considered appropriate by NTP or FDA for common tumors because it can result in an inappropriately high number of false positive results.

Hepatocellular adenoma and carcinoma combined

Table 1 (below) illustrates the impact of these choices on the determination of statistical significance of the combined hepatocellular adenoma and carcinoma in high dose female mice compared to the control group. The only combination of choices that results in a statistically significant increase (p=0.049) is Option 7 (highlighted), which is the only result presented in the HID. Of note, none of the other seven options results in a statistically significant increase in liver tumors. For example, if the statistical criteria employed by NTP (e.g., p<0.01, total number of mice) is used, there is no statistically significant increase in liver tumors among the high dose female mice.

Table 1. Results of various pairwise comparisons using Fisher exact test on the incidence of hepatocellular adenoma and carcinoma combined between the high dose female mice and controls in the Amo and Matsuyama (1985) study

Option	No. of mice	One- or two-sided test?	p -value for significance?	Calculated p-value	Statistically significant?
1	Total[a]	One	p<0.01[c]	0.064	NO
2	Total	Two	p<0.01	0.097	NO
3	Effective[b]	One	p<0.01	0.049	NO
4	Effective	Two	p<0.01	0.092	NO
5	Total	One	p<0.05[d]	0.064	NO
6	Total	Two	p<0.05	0.097	NO
7	Effective	One	p<0.05	0.049	YES
8	Effective	Two	p<0.05	0.092	NO

[6] Amo, H., Matsuyama, M., 1985. Subchronic and chronic effects of feeding of large amounts of acetaminophen in B6C3F1 mice. Nihon Eiseigaku Zasshi 40, 567-574. (See page 570)

[7] Haseman J (1984) Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Env Hlth Perspect 58:385-92.

[8] OECD (2010) Draft Guidance Document 116 – Section 4, p 25-26 http://www.oecd.org/chemicalsafety/testing/45335229.pdf

1625 Eye Street, NW, Suite 600 | Washington, DC 20006 | T 202.429.9260 | F 202.223.6835 | chpa.org | @chpa

[a] Total = the number of animals assigned to the study
[b] Effective = the number of animals alive at the occurrence of the first tumor at any site
[c] Statistically significant at p is less than 0.01 (common tumors)
[d] Statistically significant at p is less than 0.05 (uncommon tumors)

Amo and Matsuyama (1985) reported that there was no statistically significant increase in liver tumors, but the publication did not provide the details of their statistical methodology. However, it is apparent that they chose an approach other than Option 7 in Table 1. It is possible that the study authors chose to use a two-sided test since they noted that the rate of tumor-bearing females was lower in both dose groups compared to the controls. It is possible that the study authors used a p-value of 0.01 since they noted that liver tumors occur at a high spontaneous rate in mice. And, it is possible that the study authors used the total number of mice assigned to the study, not the number alive at the occurrence of the first tumor at any site. If they had made any one of these three choices, the result would have been that there is no statistically significant increase, which is exactly what they reported.

Pituitary gland adenomas

Table 2 shows that the statistical evaluation of the pituitary gland adenomas among the high dose females compared to the controls gives results similar to those for liver tumors. Options 5 and 7 (highlighted) show a statistically significant increase in pituitary gland adenomas. The remaining six options did not demonstrate a statistically significant increase. Pituitary gland adenomas are common tumors in female B6C3F1 mice. For example, in the NTP cancer bioassay of acetaminophen, the incidence of pituitary gland adenomas in female B6C3F1 mice was 30%
(14/46) among the control group. When the appropriate p-value of 0.01 is used as the criterion for statistical significance for this common tumor or when a two-tailed test is used, none of the pairwise comparisons is statistically significant.

Table 2. Results of various pairwise comparisons using Fisher exact test on the incidence of pituitary gland adenomas between the high dose female mice and controls in the Amo and Matsuyama (1985) study

Option	No. of mice	One- or two-sided test?	p-value for significance?	Calculated p-value	Statistically significant?
1	Total[a]	One	p<0.01[c]	0.038	NO
2	Total	Two	p<0.01	0.055	NO
3	Effective[b]	One	p<0.01	0.028	NO
4	Effective	Two	p<0.01	0.051	NO
5	Total	One	p<0.05[d]	0.038	YES
6	Total	Two	p<0.05	0.055	NO
7	Effective	One	p<0.05	0.028	YES
8	Effective	Two	p<0.05	0.051	NO

[a] Total = the number of animals assigned to the study
[b] Effective = the number of animals alive at the occurrence of the first tumor at any site
[c] Statistically significant at p < 0.01 (common tumors)
[d] Statistically significant at p < 0.05 (uncommon tumors)

1625 Eye Street, NW, Suite 600 | Washington, DC 20006 | T 202.429.9260 | F 202.223.6835 | chpa.org | @chpa

Comparison of the tumor results of Amo and Matsuyama (1985) with the NTP (1993) cancer bioassay

Since Amo and Matsuyama (1985) and the NTP (1993) cancer bioassay were both conducted using B6C3F1 mice given acetaminophen at the same high concentrations in the diet, it is instructive to compare the liver tumor results in the two studies. In male mice, both studies showed ***decreases*** in the incidences of liver tumors, expressed as adenomas and carcinomas combined, at the high dose (Table 3)[9]. Amo and Matsuyama (1985) did not report any statistical results for the decrease in liver tumors in high dose male mice. NTP reported a p-value of 0.028 for the reduced incidence of liver tumors among the high dose male mice compared to controls; however, using NTP's statistical criterion of $p<0.01$, this reduction is not statistically significant.

Table 3. Comparison of the incidence of liver adenoma and carcinoma combined among male and female B6C3F1 mice in the Amo and Matsuyama (1985) and NTP (1993) carcinogenicity studies.

Study	Concentration of acetaminophen in the diet			
	0 ppm	600 ppm	3000 ppm	6000 ppm
Amo and Matsuyama (1985)				
Male mice	13/43 (30%)	Not determined	12/39 (31%)	6/45 (13%)
Female mice	2/49 (4%)	Not determined	2/46 (4%)	8/50 (16%)
NTP (1993)				
Male mice	16/50 (32%)	9/50 (18%)	10/50 (20%)	7/50 (14%)
Female mice	3/49 (6%)	4/50 (8%)	7/50 (14%)	3/49 (6%)

In female mice, the increase in the liver tumors observed at the high dose in the Amo and Matsuyama (1985) study was not confirmed in the NTP (1993) cancer bioassay, i.e., the incidence of liver tumors was virtually the same in the control and high dose groups (Table 3). In short, there is no evidence that acetaminophen causes liver tumors in male or female B6C3F1 mice in the NTP cancer bioassay.

Table 4 compares the results of the pituitary gland adenoma data in B6C3F1 mice in the Amo and Matsuyama (1985) and NTP (1993) carcinogenicity studies.[10] In the NTP (1993) cancer bioassay, there was not even suggestive evidence of an increase in pituitary gland adenomas in either male or female mice given the same high dose level as the mice in the Amo and Matsuyama (1985) study.

[9] Table 3 appears in our previous comments as Table 25. It is re-printed here for the convenience of the CIC members.

[10] Table 4 appears in our previous comments as Table 26. It is re-printed here for the convenience of the CIC members.

1625 Eye Street, NW, Suite 600 | Washington, DC 20006 | T 202.429.9260 | F 202.223.6835 | chpa.org | @chpa

Table 4. Comparison of the incidence of pituitary gland adenomas among male and female B6C3F1 mice in the Amo and Matsuyama (1985) and NTP (1993) carcinogenicity studies

Study	Concentration of acetaminophen in the diet			
	0 ppm	600 ppm	3000 ppm	6000 ppm
Amo and Matsuyama (1985)				
Male mice	0/43 (0%)	Not determined	1/39 (3%)	1/45 (2%)
Female mice	2/49 (4%)	Not determined	3/46 (7%)	9/50 (18%)
NTP (1993)				
Male mice	0/48 (0%)	0/39 (0%)	0/39 (0%)	0/46 (0%)
Female mice	14/46 (30%)	16/43 (37%)	7/42 (17%)	14/45 (31%)

Conclusions

When the results of the Amo and Matsuyama (1985) and NTP (1993) are considered either individually or collectively, the overwhelming weight of evidence indicates that acetaminophen did not cause an increase in any tumor in B6C3F1 mice of either sex. Amo and Matsuyama
(1985) reported no statistical difference in the incidences of any tumor in male or female B6C3F1 mice. Although the HID re-analyzed the data from this study and reported statistically significant increases in liver tumors and pituitary gland adenomas among the high dose female mice, these increases are not statistically significant using the appropriate statistical methods used by the NTP. Importantly, there is not even a hint of an increase in liver tumors or pituitary gland adenomas in female mice of the same strain given the same dose levels of acetaminophen in the NTP bioassay. The findings of the NTP (1993) cancer bioassay support the conclusions of Amo and Matsuyama (1985) that acetaminophen poses "no carcinogenic hazard for B6C3F1 mice."[11]

It is misleading to re-interpret the results of studies without disclosing that the interpretation of the results presented is not that of the study authors. The HID presented OEHHA's re-interpretation of the Amo and Matsuyama results at least 6 times and the study authors' conclusions on those same tumors zero times. In the future, when the results of a carcinogenicity study are re-analyzed in a HID, we urge that it be made transparent. The conclusions of the study authors should be presented along with the conclusions of the re-analysis. Every effort should be made to describe why the conclusions of the study authors and the re-analysis differ.

[11] Amo, H., Matsuyama, M., 1985. Subchronic and chronic effects of feeding of large amounts of acetaminophen in B6C3F1 mice. Nihon Eiseigaku Zasshi 40, 567-574.

1625 Eye Street, NW, Suite 600 | Washington, DC 20006 | T 202.429.9260 | F 202.223.6835 | chpa.org | @chpa

CHPA respectfully submits these supplemental comments. We hope this submission will be helpful to the CIC members as they prepare for the upcoming meeting.

Sincerely,

Barbara A. Kochanowski, Ph.D.
Senior Vice President, Regulatory & Scientific Affairs

1625 Eye Street, NW, Suite 600 | Washington, DC 20006 | T 202.429.9260 | F 202.223.6835 | chpa.org | @chpa