# Memorandum

Date: 02/28/07

From: T. Scott Thurmond, Ph.D.

Subject: Review of available data from the Ramazzini Foundation on their end-of-life aspartame study in Sprague Dawley rats

To: David H. Hattan, Ph.D.

## I. INTRODUCTION

This memorandum represents a review of the data submitted at the request of the FDA by the European Ramazzini Foundation (ERF) for their end-of-life aspartame (ASP) feeding study in Sprague Dawley rats. The review will be based upon information contained in a study narrative submitted by the ERF, and data summary tables and charts included in the package.

The published results from this study run counter to the Agency's conclusions from evaluations of prior ASP studies, and we requested more detailed study data to determine whether we should reassess our previous position regarding the safety of this compound. The FDA asked that ERF submit for review a copy of the study protocol, the study summary tables, and all of the individual animal data including body weights, food intake, date of death, clinical observations, and hematology data. We also requested individual gross and microscopic pathology data for all study animals, and the diagnostic criteria used to classify/grade the pathological changes for all reported major significant changes, as well as all available historical control data from the ERF laboratory for the rat strain used. In their subsequent response the only individual animal data submitted were for the dates of death and organ-related occurrence of neoplasia; however, the criteria used to classify these tumors were not included.

## II. IDENTIFICATION OF STUDY

The study was titled, "Long-term carcinogenicity bioassay to evaluate the potential biological effects, in particular carcinogenic, of aspartame administered in feed to Sprague-Dawley rats." It was conducted by the European Foundation of Oncology and Environmental Sciences, "B. Ramazzini," Via Guerrazzi, 18, 40124 Bologna, Italy, and the Cesare Maltoni Cancer Research Center (CMCRC) of the European Ramazzini Foundation, Castello di Bentivoglio (Bologna), Via Saliceto, 3, 40010 Bentivoglio, Bologna, Italy, in their facilities.

## III. GOOD LABORATORY PRACTICES

The study narrative states that Good Laboratory Practices (GLPs) were followed. **It is noted, however, that no signed GLP or QA statement or other evidence of GLP compliance was included with the submission.**

## IV. EXECUTIVE SUMMARY

No Executive Summary was included in the study narrative; however, the paper published by the ERF in *Environmental Health Perspectives* on their findings did contain an abstract of the study results and conclusions (see reference 1).

## V. MATERIALS AND METHODS

### A. Test substance

Food grade ASP (greater than 98% purity) was produced by Nutrasweet and supplied to ERF by Giusto Faravelli S.P.A of Milan, Italy. Infrared absorption spectrophotometry was used to determine purity. The study narrative states that at the start of the study the test substance was analyzed for proper concentration, and the stability of ASP in the feed was evaluated **(these data were not included in the data submission)**.

### B. Test substance as administered

ASP was administered in the feed at concentrations of 100,000, 50,000, 10,000, 2,000, 400, and 80 ppm.

### C. Animal feed and water

Animals were fed *ad libitum* a pelleted diet used by ERF for over 30 years ("Corticella type," Laboratorio Dottori Piccioni). Analyses were conducted for nutritional value and contaminants **(no data for these analyses were included)**. Feed was used within two months of production. Tests for stability of ASP in the feed were conducted only at the start of the study, **although no results were included in the narrative.**

Drinking water was analyzed "periodically" for the possible presence of "pollutants" **(these pollutants were not identified)**. All values for the analytes were consistently reported to be within acceptable ranges **(no data for these analyses were included)**.

### D. Test animals

Male and female Sprague-Dawley rats from an in-house colony were used for this study. The colony has been utilized for various ERF experiments for over 30 years. Information on the original source of the colony was not provided. Animals were ear-punched for identification, and randomized into treatment groups. All animals were housed by sex, 5 to a cage in makrolon cages (41 X 25 X 15 cm), with stainless steel tops and a shallow layer of white wood shavings as bedding **[no indication that the bedding was autoclaved, or otherwise decontaminated]**. **[It should be noted that cage size used for this study would give approximately 50% less space than that recommended in the Decreto Legislativo 116 1992 (the Italian law cited by the authors as regulating the use of animals for scientific purposes), for rats of 500 grams or greater in weight.]**

### E. Experimental design

The study design was as follows: The test compound was administered in the feed to all treatment groups *ad libitum* from 8 weeks of age to spontaneous death. The control groups were fed the basal diet for the duration of their lifetime.

Group I: 200 rats (100 M and 100 F), 100,000 ppm.
Group II: 200 rats (100 M and 100 F), 50,000 ppm.
Group III: 200 rats (100 M and 100 F), 10,000 ppm.
Group IV: 300 rats (150 M and 150 F), 2,000 ppm.

Group V: 300 rats (150 M and 150 F), 400 ppm.
Group VI: 300 rats (150 M and 150 F), 80 ppm.
Group VII: 300 rats (150 M and 150 F), 0 ppm.

Total animals = 1800

Reviewer's comment: The selection of the dose range was made based upon the "assumed daily intake by humans of 5,000, 2,500, 500, 100, 20, 4 or 0 mg/kg b.w.," with no further explanation. No reference for a source for these numbers was included. It should be noted that the FDA Acceptable Daily Intake (ADI) at the 90$^{th}$ percentile level for aspartame is 50 mg/kg b.w./day, although market research has shown that the actual intake per person per day in the US averages between 6% and 10% of the ADI (Market Research Corporation of America survey between July 1991 to June 1992).

## F. Experimental measurements

### 1. Water and feed consumption

Mean water and feed consumption per cage was measured for the first 100 animals in each group (50 M and 50 F) once a week for the first 13 weeks and every two weeks thereafter until 110 weeks of age.

Reviewer's comment: There is no indication that water and feed consumption was measured for the same groups of animals throughout the study. There is also no explanation of the method used to select the first 100 animals in each group; most studies of this type will assess effects in all animals on study.

### 2. Body Weights

Individual body weights for all animals were measured weekly for the first 13 weeks of study, and every two weeks thereafter until 110 weeks of age, and then every 8 weeks until study's end.

### 3. Clinical observations

Clinical evaluations for the presence of gross neoplasms were conducted every week for the first 13 weeks, and then every two weeks until the end of the study. Animals were evaluated for overall health status and behavior three times a day, and clinically for gross changes every two weeks. **[There is no indication that all animals were assessed, or what behaviors were observed.]**

### 4. Pathology

Animals found dead were kept at 4°C and necropsied within 16 hours following death. All necropsies were either performed by, or supervised by a qualified pathologist. A complete list of the 36 organs and tissues taken for histopathologic analyses is shown on pages 11 and 12 of the study narrative.

All organs and tissues, with the exception of bone, were fixed in 70% ethanol. Bone samples were fixed in 10% formalin and decalcified in 10% formaldehyde + 20% formic acid in an

aqueous solution. Tissues were routinely processed and 3 – 5 μm sections mounted on glass slides. Tissue slides were stained in hematoxylin and eosin, with S100 staining used to further characterize schwannomas and chromagranin A staining used to further characterize olfactory neuroblastomas.

Slides were reviewed by a "junior pathologist" and then by a "senior pathologist." An NIEHS pathology working group (PWG) provided a second opinion for a set of malignant lesions and their precursors.

Reviewer's comment: There is no indication that an effort was made to reduce potential observational bias on the part of the primary reviewers in evaluating the histopathology slides. Although the ERF submitted histopathology slides to the NIEHS PWG for a second opinion, the small number (60) did not represent an adequate cross section of the total number of slides (~36,000) to reduce this possible bias.

### G. Statistics

Two statistical tests were used to analyze neoplastic and non-neoplastic lesion incidence data. The Cochran-Armitage trend test was used to test for linear trends in tumor incidence. Also used was the poly K-test, a survival adjusted quantal response modification of the Cochran-Armitage test that takes into account survival. **[There is no mention in the narrative of statistical testing on non-tumorous lesions or on in-life variables, e.g., body weight, feed consumption, etc.]**

The National Toxicology Program (NIEHS) conducted the statistical evaluation of the tumor data using software developed by it for analysis of two-year carcinogenicity studies. There is not sufficient information regarding this software to assess the appropriateness of using it to analyze data from an end-of-life study.

## VI. RESULTS

### A. Clinical observations

The only clinical observations included in the study narrative was a general comment on the lack of behavioral changes in ASP treated animals and an observed yellowing of the fur of the high dose animals.

Reviewer's comments

The authors relate this change in coat color to the formaldehyde produced as a breakdown product of ASP in the rat. This reviewer questions this explanation because the half-life of formaldehyde in blood has been shown to be short (1 – 2 minutes, reference 2). The European Food Safety Authority (EFSA) conducted a thorough evaluation of this hypothesized occurrence and included their conclusions in their overall evaluation of this study (reference 3)

### B. Water consumption

Summary data on group mean water consumption for male and female animals is presented in Tables 2 and 3, and Figures 1 and 2 in the study narrative. The study narrative states that no differences were

observed among the various groups, apart from a slight decrease in consumption from 72 weeks of age in males and females treated at the highest concentration of ASP.

Reviewer's comments: The data show only the mean values for all the animals of a given sex within each group. There are no standard deviations or standard errors included in the data, and no statistical evaluations were conducted. The observations made in the study narrative cannot be confirmed because individual animal data were not provided.

**C. Feed consumption**

Summarized data on feed consumption are shown in Tables 4 and 5, and Figures 3 and 4 in the study narrative. The narrative states that a dose-related difference in feed consumption was observed among the dose and control groups for males and females.

Reviewer's comments: Only group mean values are shown in the data, with no standard errors or deviations. There is an apparent correlation between increased dose and decreased feed consumption throughout the study. In previous aspartame studies we have observed similar high dose-related decreases in feed consumption that were attributable to taste aversion; however, the authors of the present study report do not address this issue. There were no intergroup statistical evaluations conducted to establish the significance of the feed consumption differences. Data on mean feed consumption at 105 weeks of age (97 weeks on study) are shown in Table 1.

**Table 1.** Study animal mean feed consumption (grams) at 105 weeks of age (97 weeks on study) by dose group (ppm)

| Sex | 100,000 | 50,000 | 10,000 | 2,000 | 400 | 80 | 0 |
|---|---|---|---|---|---|---|---|
| Male | 19.2 | 21.6 | 22.4 | 22.2 | 22.8 | 21.3 | 21.4 |
| Female | 15.8 | 17.0 | 18.8 | 18.7 | 22.9 | 21.2 | 20.5 |

**D. Body weights**

Summarized data on body weight changes are shown in Tables 6 and 7, and Figures 5 and 6 in the study narrative. The narrative states that no significant differences in mean body weights were observed between treatment and control groups for either sex, with the exception of decreases for both sexes in the 100,000 ppm group.

Reviewer's comments: No statistics were reported for this variable. Only group mean values are shown in the data, with no standard errors or deviations. The decreased weight gain noted in the narrative for the 100,000 ppm male and female groups could be attributable to the decreased food intake, as well as decreased caloric density, for these animals, although this was not discussed.

The narrative's authors also do not address the disparities evident in body weight change and feed consumption during the study. For instance, at 105 weeks (97 weeks on study) the high dose females had a marked decrease in feed consumption (~23%, Table 1) but a weight decrease of only 4.5% versus the controls (Table 2). One would expect to see a more significant weight decrease given the high drop off in feed consumption. In previous studies on aspartame submitted to the Agency a much stronger correlation was observed between decreased feed consumption in the high dose treatment groups and

weight loss. Data on mean body weights at 105 weeks of age (97 weeks on study) are shown in Table 2.

**Table 2.** Study animal mean body weights (grams) at 105 weeks of age (97 weeks on study) by dose group (ppm)

| Sex | 100,000 | 50,000 | 10,000 | 2,000 | 400 | 80 | 0 |
|---|---|---|---|---|---|---|---|
| Male | 498.8 | 516.5 | 509.6 | 513.5 | 491.3 | 514.8 | 509.5 |
| Female | 338.4 | 365.2 | 360.2 | 365.3 | 369.0 | 350.4 | 354.4 |

## E. Survival

Summarized data on survival are shown in Tables 8 and 9, and Figures 7 and 8 in the study narrative. No "substantial" differences in survival were reported for any of the groups with the exception of the control groups, which had slightly decreased survival starting at 104 weeks of age.

Reviewer's comments: Data from figure 8 (female survival) appear to indicate that the two highest ASP dose groups (50,000 and 100,000 ppm) had the greatest percentage of animals surviving from about 88 weeks of age until the end of the study. No discussions of these data are included in the narrative. Overall survival data for study animals at 105 weeks of age (97 weeks on study) are shown in Table 3.

**Table 3.** Study animal percent survival at 105 weeks of age (97 weeks on study) by dose group (ppm).

| Sex | 100,000 | 50,000 | 10,000 | 2,000 | 400 | 80 | 0 |
|---|---|---|---|---|---|---|---|
| Male | 40.0 | 42.0 | 34.0 | 36.7 | 32.0 | 38.7 | 38.0 |
| Female | 55.0 | 49.0 | 47.0 | 42.0 | 41.3 | 39.3 | 39.3 |

Although individual data on survival are included in the appendices, no statistical assessment was apparently conducted on these data. The narrative also does not provide historical data on survival for animals from this in-house colony.

## F. Non-cancerous lesions

Inflammatory lesions observed in a variety of organs and tissues are summarized in Table 10 in the narrative. The authors state that acute and chronic inflammatory processes, particularly in the lungs and kidneys, were the most commonly observed lesions. A high incidence of bronchopneumonia was observed in animals of both sexes in treated and control groups. The study authors postulate that this condition may have been related to the spontaneous death of the animals.

Reviewer's comment:

The authors do not further address the high incidence of bronchopneumonia other than to postulate it may be the cause of "spontaneous death" in all groups. This reviewer notes that inflammatory pulmonary lesions are not commonly reported in aging rats unless a concomitant infectious process is also present (reference 4) See attachment 1 for complete review of the pathology findings for this study.

## G. Cancerous lesions

The occurrences of benign and malignant tumors are summarized in tables 11 and 12 of the narrative, respectively. Individual animal tumor data are shown in Appendix I of the narrative and statistical analyses are shown in Appendix II. The authors statistically analyzed all total benign and total malignant tumors and reported that:

- Females with benign tumors showed a significant positive trend **(this reviewer assumes that here they're relating this change to increasing dose)**, while males that were dosed at 10,000 ppm only had a significant increase in benign tumors compared to controls.
- Malignant tumor-bearing animals occurred with a significant positive trend in males and in females, and in a pair-wise comparison females at 50,000 ppm were the only animals that had a significantly greater burden of malignant tumors than did the controls.

The authors analyzed the occurrence of several individual tumor types,

- preneoplastic and neoplastic lesions of the olfactory epithelium (Table 15 of the narrative)
- preneoplastic and neoplastic lesions of the renal pelvis and ureter (tables 16 – 18 of the narrative)
- preneoplastic and neoplastic lesions of the bladder (Table 19 of the narrative)
- malignant tumors of the brain (Table 20 of the narrative)
- malignant schwannomas of peripheral nerves (Table 21 of the narrative)
- Lymphomas and leukemias (Table 22 of the narrative) **[these tumor types were lumped together in their statistical analyses because the authors considered that the "distinction between them is artificial." They cited a reference by Harris *et al* (2001) in support of their decision. (see narrative for complete reference)]**

In their evaluation of these tumor types the authors noted:

- A significant positive trend exists for olfactory hyperplasia for both males and females
- A dose-related increase in the incidence of dysplastic hyperplasia and dysplastic papillomas of the transitional epithelium of the renal pelvis and the ureter in females. They also reported a positive trend in the occurrence of renal carcinomas for females ($p \leq .05$), with a significant increase observed in the high dose females ($p \leq .05$) **[It should be noted that the Pathology Working Group of NTP agreed with the diagnosis of hyperplasia of the renal pelvic transitional epithelium, but did not support the "dysplastic" descriptor for this diagnosis].**
- A positive trend in males for the incidence of malignant schwannomas of the peripheral nerves ($p \leq .05$)
- A positive trend in the incidence of lymphomas and leukemias in both males and females ($p \leq .05$), with significance versus the controls for ASP doses of 100,000 ppm, 50,000 ppm, 10,000 ppm, 2,000 ppm and 400 ppm for the females.

Tables 4, 5 and 6 show the incidences of these tumor types by sex and dose group.

**Table 4.** Incidence by sex and ASP dose level of major neoplastic types reported by the authors (number animals diagnosed with neoplasms/number of animals examined).

| Neoplasm Type | Olfactory hyperplasia | | | | Malignant schwannomas | | | |
|---|---|---|---|---|---|---|---|---|
| Sex | M | F | Percent of total | | M | F | Percent of total | |
| Dose Group | | | M | F | | | M | F |
| Grp I (100,000 ppm) | 14/100 | 19/100* | 14 | 19 | 4/100 | 2/100 | 4 | 2 |
| Grp. II (50,000 ppm) | 12/100 | 21/100 | 12 | 21 | 3/100 | 1/100 | 3 | 1 |
| Grp. III (10,000 ppm) | 7/100 | 17/100 | 7 | 17 | 2/100 | 1/100 | 2 | 1 |
| Grp. IV (2,000 ppm) | 4/150 | 13/150 | 3 | 9 | 2/150 | 3/150 | 1 | 2 |
| Grp. V (400 ppm) | 9/150 | 11/150 | 6 | 7 | 3/150 | 0/150 | 2 | 0 |
| Grp. VI (80 ppm) | 3/150 | 5/150 | 2 | 3 | 1/150 | 2/150 | 0.7 | 1 |
| Grp. VII (0 ppm) | 1/150 | 6/150 | 0.7 | 4 | 1/150 | 0/150 | 0.7 | 0 |

*One hyperplasia with atypia

**Table 5.** Incidence reported by the authors of dysplastic lesions and carcinomas of the transitional cell epithelium of the renal pelvis and ureter (number of animals diagnosed with neoplasms/number of animals examined).

| Renal Pelvis Neoplasm Type | Dysplastic hyperplasia | | Dysplastic Papilloma | | Carcinoma | |
|---|---|---|---|---|---|---|
| Sex | M | F | M | F | M | F |
| Dose Group | | | | | | |
| Grp I (100,000 ppm) | 3/100 | 8/100 | 0/100 | 3/100 | 1/100 | 4/100 |
| Grp. II (50,000 ppm) | 2/100 | 6/99 | 0/100 | 1/99 | 1/100 | 3/99 |
| Grp. III (10,000 ppm) | 2/100 | 6/100 | 0/100 | 1/100 | 1/100 | 3/100 |
| Grp. IV (2,000 ppm) | 4/150 | 6/150 | 0/150 | 1/150 | 1/150 | 3/150 |
| Grp. V (400 ppm) | 4/149 | 5/150 | 1/149 | 1/150 | 0/149 | 3/150 |
| Grp. VI (80 ppm) | 3/149 | 4/150 | 0/149 | 1/150 | 0/149 | 1/150 |
| Grp. VII (0 ppm) | 1/150 | 2/150 | 0/150 | 0/150 | 0/150 | 0/150 |

**Table 6.** Incidence reported by the authors for lymphomas and leukemias (number of animals diagnosed versus the number of animals examined).

| Lymphomas-Leukemias | | | | | | |
|---|---|---|---|---|---|---|
| | Total with Diagnoses/Total animals (%) | | Lymphoblastic Lymphoma (total diagnosed) | | Lymphoblastic Leukemia (total diagnosed) | |
| Dose Group | M | F | M | F | M | F |
| Grp I (100,000 ppm) | 29/100 (29) | 25/100 (25) | 0 | 1 | 0 | 0 |
| Grp. II (50,000 ppm) | 20/100 (20) | 25/100 (25) | 0 | 2 | 0 | 0 |
| Grp. III (10,000 ppm) | 15/100 (15) | 19/100 (19) | 0 | 2 | 1 | 0 |
| Grp. IV (2,000 ppm) | 33/150 (22) | 28/150 (18.7) | 0 | 5 | 0 | 1 |
| Grp. V (400 ppm) | 25/150 (16.7) | 30/150 (20) | 0 | 7 | 0 | 0 |
| Grp. VI (80 ppm) | 23/150 (15.3) | 22/150 (14.7) | 3 | 3 | 0 | 0 |
| Grp. VII (0 ppm) | 31/150 (20.7) | 13/150 (8.7) | 0 | 2 | 1 | 0 |

**Table 6 (cont'd)**

| Lymphomas-Leukemias | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Lymphocytic Lymphoma (total diagnosed) | | Lymphoimmuno-blastic lymphoma (total diagnosed) | | Histiocytic sarcoma/monocytic leukemia (total diagnosed) | | Myeloid leukemia (total diagnosed) | |
| Dose Group | M | F | M | F | M | F | M | F |
| Grp I (100,000 ppm) *100 animals/sex* | 0 | 2 | 24 | 11 | 4 | 9 | 1 | 2 |
| Grp. II (50,000 ppm) *100 animals/sex* | 0 | 0 | 13 | 10 | 7 | 12 | 0 | 1 |
| Grp. III (10,000 ppm) *100 animals/sex* | 0 | 2 | 8 | 3 | 5 | 12 | 1 | 0 |
| Grp. IV (2,000 ppm) *150 animals/sex* | 1 | 1 | 15 | 8 | 17 | 12 | 0 | 1 |
| Grp. V (400 ppm) *150 animals/sex* | 0 | 2 | 22 | 8 | 3 | 14 | 0 | 0 |
| Grp. VI (80 ppm) *150 animals/sex* | 0 | 5 | 12 | 6 | 6 | 8 | 2 | 0 |
| Grp. VII (0 ppm) *150 animals/sex* | 0 | 2 | 19 | 5 | 10 | 4 | 1 | 0 |

Reviewer's comments

See attachment 1 for review by CFSAN Pathology

### H. Statistics

See attachment 2 for review by CFSAN Statistics

## VII. DISCUSSION

A complete and acceptably rigorous review of this study could not be performed due to the lack of critical individual animal data (e.g., individual weights, clinical observations, etc.), and the use of mean values for other data without the inclusion of standard deviations or standard errors. A lack of current historical control data for the animal colony also makes the author's clinical observations and pathologic findings difficult to verify.

The end-of-life study design creates problems in that the increase in background pathology over the total life of the animal can confound interpretation of changes that may be related to treatment (Federal Register, March 14, 1985, pp. 10371-10442; Office of Science and Technology Policy, Part II, Chemical Carcinogens; A review of the science and its associated principles). This increase in background pathology is one reason that most official guidelines suggest a testing duration for rat carcinogenicity studies of two years, which constitutes the major portion of their life span. Another issue with the design of this study is the inordinately wide ASP dose range (six dose groups from 80 to 100,000 ppm, 4 to 5,000 mg/kg b.w./day), which is purportedly based on an "assumed" range of human daily intake (they do not reference a source for their numbers). There is no other rationale for the selection of these doses mentioned in the narrative, or in the two publications relating to this study by the study's authors. Actual reported intakes by consumers are in the range of 3 -- 5 mg/kg/day.

The study narrative discusses in-life variables such as weight gain, feed and water consumption and survival; however, there were no statistical analyses assessing intragroup changes in these variables. The observed decreased weight gain in the high dose ASP groups relative to the other treatment and the control groups may be attributable to taste aversion, which has been observed in high dose groups in other ASP studies. The survival data are of particular interest in that the two high dose groups for both sexes had a higher percentage of surviving animals at 105 weeks of age than did the lower dose groups and the control group. Whether this is related to the overall lower weight gain in these groups is difficult to determine. However, numerous studies have shown that dietary restriction in rodents can lead to increased longevity, as well as altered incidences of cancer over time (reference 5). There was no discussion of any of these potentially important changes in the report narrative.

## VIII. SUMMARY

### General

- There is insufficient information included in the review package to allow a complete and definitive review of the ERF end-of-life rat ASP study.
- Issues such as an overly wide dose range, potential for increased background pathology due to extended duration of study, possible presence of epizootic infection among the test animals (see attachment 1), apparent taste aversion in high dose groups with concomitant effects on feed consumption and weight gain (and possibly on longevity), disparities in feed consumption and body weight change, and housing issues (i.e., small cage size) that may have increased stress on test animals, as well as unrecognized confounding factors in the study design (see attachment 2), raise questions concerning how much these study design shortcomings and uncontrolled variables may have adversely affected the outcomes of the study and the authors' objective interpretation of their results.

**Pathology (see attachment 1 for complete discussion)**

- Aspartame related findings, proposed by the study's authors are just not evident from the data presented. The study design, the data presentation and the use of diagnostic criteria are not consistent with current recommendations for the conduct of carcinogenicity studies. Specifically, the study duration over the lifetime of the test animals, the high incidence of inflammatory lesions and combining of incidences of unrelated changes in the summary tables make the reported study results highly questionable.
- Based on review of the ERF aspartame study, the pathologic changes were incidental and appeared spontaneously in these rats, which lived up to a year longer than in routine carcinogenicity studies. None of the histopathologic changes are related to treatment with aspartame.

In regard to evaluation of selected ERF slides by the NTP Pathology Working Group (PWG):

- The PWG agreed upon many of the diagnoses, however, there were also substantial differences regarding the classification of some lesions.
- It appeared that there were many slides with autolytic changes; autolytic changes impaired the PWG's ability to classify certain lesions in more detail.
- In the case of kidney lesions, PWG reported that there were inflammatory lesions associated with the proliferative changes. This certainly raises questions about the diagnostic accuracy and the significance of the reported renal changes

**Statistics (see attachment 2 for complete discussion)**

- There is no mention of blinding in necropsies and histological evaluations. To prevent conscious or unconscious bias in these evaluations, examiners should always be blinded to dose group membership when possible. Blinding is critical in ensuring the validity of experimental results.
- The rats had a very high incidence of bronchopneumonia. More importantly, there were substantial differences in the incidence of bronchopneumonia across dose groups. The dose-response for bronchopneumonia in each sex was that the lowest dose of aspartame had a substantially lower incidence rate than the control group, with the incidence increasing monotonically with dose until the incidence for the highest dose was nearly as high as the control group. This unusual dose-response pattern suggests that there may be unrecognized confounding factors in the design of the study that resulted in treatment groups that were not comparable.
- A large number of statistical tests were performed in the analysis of the study. When conducting many tests, one would expect numerous statistically significant results by chance alone. The authors should describe how they dealt with this testing multiplicity issue in the interpretation of the study.

## XIX. CONCLUSION

Based on a review and evaluation of the data provided by the ERF for their lifetime Sprague-Dawley rat study on aspartame it was not possible to confirm the reported results. There are significant shortcomings in the study design, conduct, reporting and interpretation. Additional insight and enhanced definition of the findings of this study would be provided by an internationally-sponsored expert pathology working group examination of a statistically representative sample of tissue slides from this study.

Considering the results from an array of studies, including five previously conducted negative chronic bioassays, a recently reported large epidemiology study with negative associations between the use of aspartame and the

occurrence of selected forms of tumors, and negative findings from a series of three transgenic mouse assays, the FDA concludes that the present regulations governing the use of aspartame as a food additive are supported by most of the available scientific evidence. The ERF study data does not provide sufficient evidence to alter the current FDA position; i.e., that there is reasonable certainty of no harm with the food additive use of aspartame.

## REFERENCES


1. Soffritti, M, *et. al.* First Experimental Demonstration of the Multipotential Carcinogenic Effects of Aspartame Administered in the Feed to Sprague-Dawley Rats. *Env Health Perspect.* 114: 379-85. 2006.
2. National Research Council, Board on Toxicology and Environmental Health, Assembly of Life Sciences, Committee on Aldehydes. 1981. *Formaldehyde and Other Aldehydes.* National Academy Press, Washington, D.C.
3. EFSA. Opinion of the Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in contact with Food (AFC) on a request from the Commission related to a new long-term carcinogenicity study on aspartame. May, 2006.
4. *Pathobiology of the Aging Rat.* vol. 1, U. Mohr, D.L. Dungworth and C.C. Capen, eds., ILSI Press, 1992.
5. Heilbronn, L and E. Ravussin. Calorie restriction and aging: review of the literature and implications for studies in humans. *Am J Clin Nutr.* 78:361-9. 2003.

Attachment 1

**DEPARTMENT OF HEALTH & HUMAN SERVICES**

Public Health Service
Food and Drug

Administration

# Memorandum

Date: December 15, 2006

From: Mathematical Statistician (Biomedical)
Division of Mathematics, HFS-705

Subject: Statistical Analysis of lifetime rat study of carcinogenicity of aspartame

To: Scott Thurmond, Ph.D.
Division of Petition Review, HFS-265

Statistical analyses were performed for a report on "Long-term Carcinogenicity Bioassays to Evaluate the Potential Biological Effects, in Particular Carcinogenic, of Aspartame Administered in Feed to Sprague-Dawley Rats" from the B. Ramazzini European Foundation of Oncology and Environmental Sciences. The study was a lifetime study of male and female rats fed aspartame in their diet. The dose groups were 0, 80, 400, 2000, 10000, 50000, 100000 ppm in diet, with 150 rats of each sex for doses 0, 80, 400, and 2000 ppm, and 100 rats of each sex for doses 10000, 50000, and 100000 ppm. Analyses of dose-response for neoplastic and non-neoplastic lesions were performed using both the Cochran-Armitage trend test (with Fisher's exact test for tests of each dose vs. control) and the poly K-test (for both trends and tests of each dose vs. control). The poly K-test adjusts for differences in survival across doses, while the Cochran-Armitage test does not.

The study differed from the typical rodent carcinogenicity study in that more animals and doses were used and the study was a lifetime study rather than the usual two-year study. Using more doses and animals gives this study better statistical power to detect dose effects than the typical study. Most rodent carcinogenicity studies are terminated at two years to control the cost and length of the study. A lifetime study may be more sensitive than a two-year study in some cases.

One problem with the report is that there is no mention of blinding in necropsies and histological evaluations. To prevent conscious or unconscious bias in these evaluations, examiners should always be blinded to dose group membership when possible. It is well-recognized that blinding is critical in ensuring the validity of experimental results. If the histological evaluations were not blinded, then an independent blinded review would be required to confirm the validity of the results in this report.

Another potential problem with the report is that the rats had a very high

incidence of bronchopneumonia. More importantly, there were substantial differences in the incidence of bronchopneumonia across dose groups, as shown in Figure 1. The dose-response for bronchopneumonia was that the lowest dose of aspartame had a substantially lower incidence rate than the control group, with the incidence increasing monotonically with dose until the incidence for the highest dose was nearly as high as the control group. This dose-response relationship was very similar for both sexes. If bronchopneumonia is spread contagiously, then standard statistical tests of incidence differences between dose groups are not valid due to the lack of statistical independence among animals. However, this unusual dose-response pattern suggests that there may be unrecognized confounding factors in the design of the study which resulted in treatment groups which were not comparable. Joiner (1981) described case studies of several experiments in which confounding factors distorted the results, including a rodent carcinogenicity study in which housing cage location was a confounding factor. Details of the conduct of the aspartame study should be reviewed for possible explanations for group differences in the incidence of bronchopneumonia.

Appendix II of the report contains statistical tests of neoplastic and non-neoplastic lesions. Because survival was similar among dose groups, the two analysis approaches (the Cochran-Armitage test and the poly K-test) gave similar results. A large number of statistical tests of aspartame effects was performed in the analysis of the study. All tests with p-values less than 0.05 were noted as statistically significant. For each of 181 kinds of lesions for each sex, a dose trend test and a test for each of the six doses vs. controls was performed, for a total of 2534 different tests of neoplastic and non-neoplastic effects of aspartame. With such a large number of tests performed, one would expect numerous statistically significant results by chance alone. The authors do not acknowledge this testing multiplicity issue or consider it in the interpretation of their results.

In Appendix II, statistically significant increases for one or both sexes for either the aspartame dose trend or for one or more aspartame dose were noted for adrenal medulla (pheochromocytoma benign, pheochromocytoma malignant), ear (squamous cell carcinoma), heart (myxoma), pancreatic islets (adenoma, combined adenoma and carcinoma), kidney (carcinoma), liver (cholagioma, hepatocellular adenoma), lung (adenoma), mammary gland (carcinoma, fibroma), oral cavity (papilloma sqamous or papilloma), ovary (cystadenoma, fibroangioma), peripheral nerve (schwannoma malignant), pituitary gland (adenoma, combined adenoma and carcinoma), spleen (fibroangioma, hemangiosarcoma), thymus (thymoma benign), uterus (polyp stromal, combined polyp stromal and sarcoma stromal), and systemic multiple organ lesions (leukemia monocytic, lymphoma lympho-immunoblastic, leukemia myeloid).
Analyses of number of animals having a particular lesion in any organ or tissue found statistically significant increases for one or both sexes for either the aspartame dose trend or for one or more aspartame dose for leukemia monocytic, lymphoma lympho-immunoblastic, leukemia myeloid, leukemia granulocytic, hemangiosarcoma, hemolymphoreticular neoplasias, histiocytic sarcoma, benign tumors, and malignant tumors.

The report concludes that "the results indicate that APM causes in our experimental conditions: 1) an increased incidence of malignant tumor-bearing animals with a positive significant trend in males ($p \leq 0.05$) and in females ($p \leq 0.01$), particularly in the females treated at 50,000 ppm ($p \leq 0.01$); 2) in females, dysplastic lesions and carcinomas of the renal pelvis and ureter combined, show a significant positive trend ($p \leq 0.01$) and a statistically significant increase in those treated at 100,000 ($p \leq 0.01$), 50,000 ($p \leq 0.01$), 10,000 ($p \leq 0.01$), 2,000 ($p \leq 0.05$) and 400 ($p \leq 0.05$) ppm; 3) a positive significant trend in both males ($p \leq 0.05$) and females ($p \leq 0.01$) and a statistically significant increase of the incidence of lymphomas-leukemias in females treated at the doses of 100,000 (pQ0.01), 50,000 (pQ0.01), 10,000 ($p \leq 0.05$), 2,000 ($p \leq 0.05$) and 400 ($p \leq 0.01$) ppm; and 4) an increased incidence of malignant schwannomas of the peripheral nerves with a positive trend in males ($p \leq 0.05$)."

**Summary**

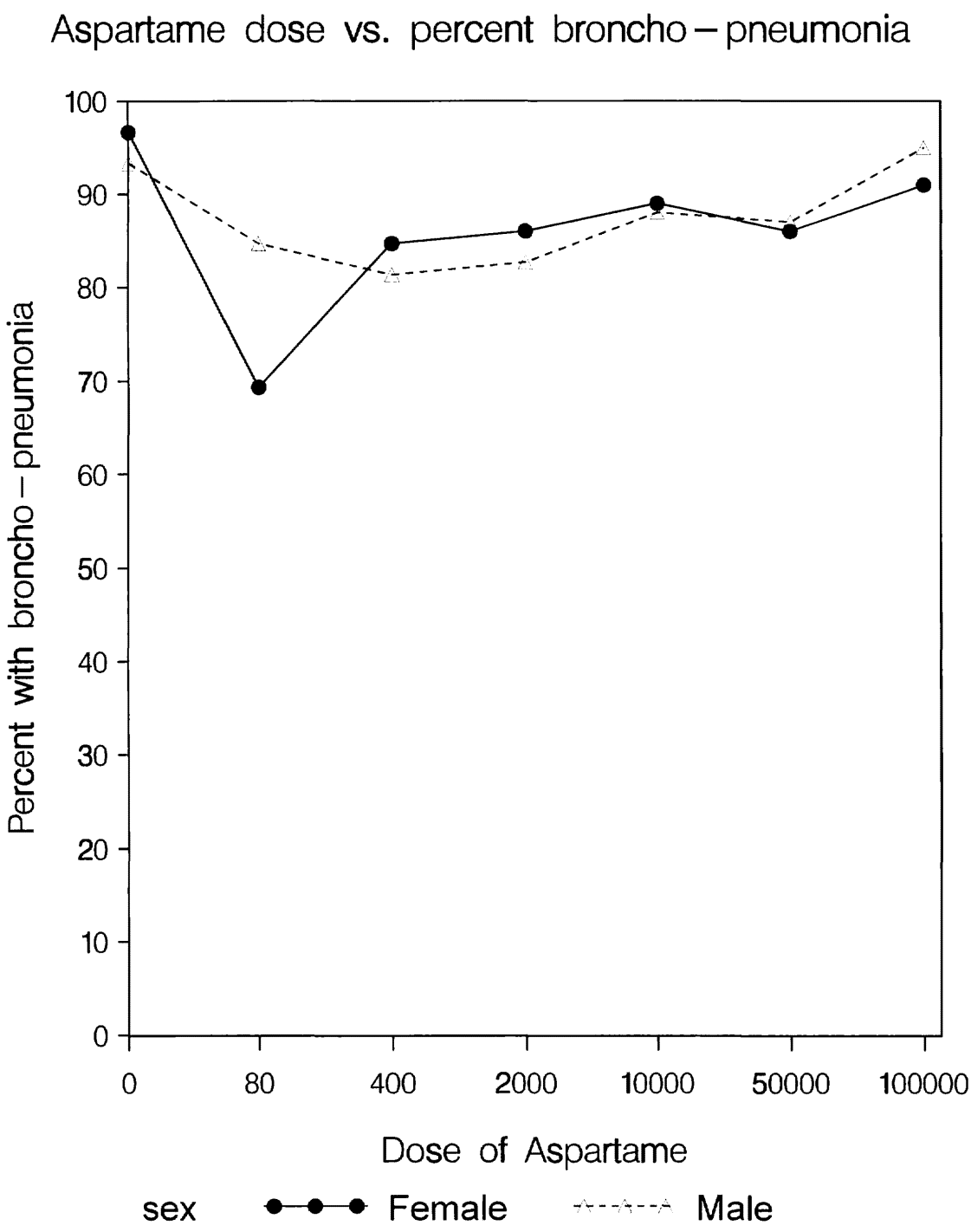
Statistical analyses were performed for a report on "Long-term Carcinogenicity Bioassays to Evaluate the Potential Biological Effects, in Particular Carcinogenic, of Aspartame Administered in Feed to Sprague-Dawley Rats". The submission reports "multipotential carcinogenic effects": carcinogenic effects at numerous sites in both sexes.

Several factors raise serious concerns about the validity of these results. First, there is no mention of blinding in necropsies and histological evaluations. To prevent conscious or unconscious bias in these evaluations, examiners should always be blinded to dose group membership when possible. Blinding is critical in ensuring the validity of experimental results. Second, the rats had a very high incidence of bronchopneumonia. More importantly, there were substantial differences in the incidence of bronchopneumonia across dose groups. The dose-response for bronchopneumonia in each sex was that the lowest dose of aspartame had a substantially lower incidence rate than the control group, with the incidence increasing monotonically with dose until the incidence for the highest dose was nearly as high as the control group. This unusual dose-response pattern suggests that there may be unrecognized confounding factors in the design of the study which resulted in treatment groups which were not comparable. Third, a large number of statistical tests were performed in the analysis of the study. With such a large number of tests performed, one would expect numerous statistically significant results by chance alone. The authors should describe how they dealt with this testing multiplicity issue in the interpretation of the study.

Curtis Barton, Ph.D.

Reference: Joiner, Brian L. Lurking Variables: Some Examples, *The American Statistician,* 1981, *35*, 227-233.

# Figure 1


Aspartame dose vs. percent broncho – pneumonia
Percent with broncho – pneumonia
100
90
80
70
60
50
40
30
20
10
0
0
80
400
2000
10000
50000
100000
Dose of Aspartame
sex
Female
Male

Attachment 2

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Food and Drug

Administration

# Memorandum

**Date:** January 08, 2007

**From:** Director, Division of General Scientific Support, OSAS, HFS-715

**Subject:** Pathology comments: Ramazzini Foundation Aspartame study

**To:** Scott Thurmond, Ph.D., Office of Food Additive Safety, HFS-265

## Reference

Long-term carcinogenicity bioassays to evaluate the potential biological effects, in particular carcinogenic, of aspartame administered in feed to Sprague-Dawley rats (Protocol No. BT 6008), 4 volumes of data marked as 1, 2, 3, and 4. Report was addressed to FDA (CFSAN) and was dated February 2006.

## Background

The above referenced report by the Ramazzini Foundation describes a long-term feeding carcinogenicity study of aspartame in Sprague-Dawley rats. Aspartame was administered at concentrations (ppm) of 100.000 to group I, 50.000 to group II, 10.000 to group III, 2.000 to group IV, 400 to group V, 80 to group VI, and 0 to group VII. The number of animals varied among groups: the higher dose groups, I through III, had 100 animals per sex per group while the lower dose groups, IV through VI and the control (VII), had 150 animals per sex per group. The animals were 8 weeks old at the start of the study and the treatment continued until spontaneous death; the last animal died at 159 weeks of age. The study results were also published in Environmental Health Perspectives, March 2006, Vol. 114, n.3, pp. 379-386.

## Review Comments

My review of the Ramazzini study focused primarily on the histopathologic findings, specifically those the authors reported as treatment-related.

Based on the study results, the study authors concluded (page 19 of volume I) that aspartame causes the following effects:

"1) an increased incidence of malignant tumor-bearing animals with a positive significant trend in males ($P \leq 0.05$) and in females ($P \leq 0.01$), particularly in the females treated at 50.000 ppm ($P \leq 0.01$);

2) in females, dysplastic (atypical) lesions and carcinomas of the renal pelvis and ureter combined, show a significant positive trend ($P \leq 0.01$) and a statistically significant increase in those treated at 100.000 ($P \leq 0.01$), 50.000 ($P \leq 0.01$), 10.000 ($P \leq 0.01$), 2.000 ($P \leq 0.05$), and 400 ($P \leq 0.05$) ppm;

3) a positive significant trend in both males ($P \leq 0.05$) and females ($P \leq 0.01$) and a statistically significant increase of the incidence of lymphomas / leukemias in females treated at the doses of 100.000 ($P \leq 0.01$), 50.000 ($P \leq 0.01$), 10.000 ($P \leq 0.05$), 2.000
($P \leq 0.05$) and 400 ($P \leq 0.01$) ppm;

4) an increase in incidence of malignant schwannoma of peripheral nerves with a positive trend in males ($P \leq 0.05$)".

Comments specifically addressing the conclusions of the study authors:

1) On increased incidence of malignant tumor-bearing animals: we do not regard the parameter of 'total-tumor bearing animals' as helpful or scientifically meaningful in the evaluation of carcinogenicity studies. Combining all tumors of different cell types and/or tissues is biologically irrelevant since it ignores compound-related, underlying carcinogenic mechanisms within a specific tissue. Therefore, the conclusion 1) above by the study authors on an increased incidence of malignant tumor-bearing animals is invalid, in my opinion.

2) On renal pelvis and ureter lesions: the renal data presentation in the report is confusing due to the use of different terms (hyperplasia, dysplastic / hyperplasia, atypical) for the same change. Both terms are used interchangeably within the study report-tables (Tables 16 & 18). An explanation on the use of the parameters was not provided and the severity scale based on which the data were generated was not discussed. In routine toxicology studies, a definition of the lesion with severity modifiers is provided to make valid comparisons.

   A Pathology Working Group (PWG) of the National Institute of Environmental Health Sciences (NIEHS) (also discussed later in my comments), reviewed selected slides from the Ramazzini study. The PWG recommended to the study authors to eliminate the terminology of 'atypical hyperplasia' all together. Furthermore, the term 'papilloma' is generally used to address a benign neoplastic change but was reported as a pre-neoplastic change in Table 16 of the study report. The renal changes are, in my opinion, not directly related to aspartame. There may be other contributing factors such as inflammatory changes noted by the PWG, which may have caused the proliferative changes of the transitional cell epithelium.

On the increased incidence of lymphomas / leukemias: the incidence reported in % in the study is as follows:

| Dose Group | Dose in ppm | Males (%) | Females (%) |
|---|---|---|---|
| I | 100,000 | 29 | 25 |
| II | 50,000 | 20 | 25 |
| III | 10,000 | 15 | 19 |
| IV | 2,000 | 22 | 19 |
| V | 400 | 17 | 20 |
| VI | 80 | 15 | 15 |
| VII | 0 | 21 | 9 |

The study authors have made their conclusion of a positive significant trend in males and females and significant increases in females based on statistical analysis. I find it difficult to concur with their conclusion. Just looking at the data of lymphomas / leukemias, the incidence appears random in both males and females and unrelated to aspartame. The differences reflect the biological variability in these types of neoplasia, especially when the study authors combined different types of lymphomas, including the commonly occurring histiocytic sarcomas.

3) On malignant schwannomas of the peripheral nerves: the reported incidence in % is given below.

| Dose Group | Dose in ppm | Males (%) malignant | Females (%) malignant |
|---|---|---|---|
| I | 100,000 | 4.0 | 2.0 |
| II | 50,000 | 3.0 | 1.0 |
| III | 10,000 | 2.0 | 1.0 |
| IV | 2,000 | 1.3 | 2.0 |
| V | 400 | 2.0 | 0 |
| VI | 80 | 0.7 | 1.3 |
| VII | 0 | 0.7 | 0 |

Again, the conclusion by the study authors on schwannomas is based on the statistical differences among groups. The distribution of malignant schwannomas within treatment groups appears random in both males and females and unrelated to treatment. Furthermore, in the results section of the study report, (page 17), the study authors compare the total the number, nine (9), of malignant schwannomas in all treatment groups to the number of schwannomas in control animals, zero (0). This not a common practice of making comparisons among groups.

Pathology Working Group (PWG) chair person's report:

The Ramazzini study report stated that 'A working group of pathologists from the National Institute of Environmental Health Sciences provided a second opinion for a set of malignant lesions and their precursors related to the APM [aspartame] treatment'. To clarify, this 'second opinion' constituted a limited review of selected slide samples and was not consistent with a routine peer review of all study data.

I reviewed the PWG report prepared by Dr. James Hailey of NIHES, dated November 30, 2004. I consider the following points made by the PWG important since they impact on the reliability of the data and conclusions presented in the Ramazzini study report.

- PWG agreed upon many of the diagnoses, however, there were also substantial differences regarding the classification of some lesions.
- It appeared that there were many slides with autolytic changes; autolytic changes impaired the PWG's ability to classify certain lesions in more detail.
- In case of kidney lesions, PWG reported that there were inflammatory lesions associated with the proliferative changes. This certainly raises questions about the diagnostic accuracy and the significance of the reported renal changes.
- It is not clear to what degree the study authors integrated the PWG comments. The PWG asked the study authors to eliminate the term 'atypical hyperplasia', which was still used in the report (Table 16).

Other findings in the study:

The study reported a number of inflammatory findings in both males and females such as: meningitis, bronchopneumonia, pleuritis, pericarditis, liver abscesses, pyelonephritis and peritonitis. These changes are not common in long term studies and question the general health status of the study animals and the potential impact on the study results. The incidences of bronchopneumonia (97%) and pyelonephritis (83%) were unusually high across groups, which may have contributed to changes that were interpreted to be treatment related by the study authors.

Over-all comments:

- In their introduction (page 2) the study authors mention that the rationale for conducting the study was that previous studies did not comply with basic requirements, specifically study duration and animal numbers. In my opinion, previously reviewed aspartame studies did meet regulatory guideline requirements for carcinogenicity testing.

- The Ramazzini study was submitted with little or no peer review. The PWG performed a very limited review of only selected slides for a very large study. The PWG questioned several diagnostic criteria used in the carcinogenicity study and it appeared that the PWG's recommendations (e.g. eliminating the use of 'atypical' hyperplasia) were not applied in the study report.

- Some of the terminology of lesions in the report suggests that the Ramazzini Lab has worked independently of the veterinary toxicologic pathology community as to the use of the standardized diagnostic terminology. For example the term 'hepatoma' was used in the report, and this term has not been in use since the early '80s. An other example is that neoplastic lesions in the rat uterus were diagnosed as adenoma, fibro-adenoma, fibro-angioma, sarcoma botryoides, and fibro-lipoma, terms that apply specifically to humans.

In summary:

Aspartame related findings, proposed by the study authors are just not evident from the data presented. The study design, the data presentation and the use of diagnostic criteria are not consistent with current recommendations for the conduct of carcinogenicity studies. Specifically, the study duration over the lifetime of the test animals, the high incidence of inflammatory lesions and combining of incidences of unrelated changes in the summary tables make the reported study results highly questionable.

Based on my review of the Ramazzini aspartame study, the pathologic changes were incidental and appeared spontaneously in rats, which lived up to a year longer than in routine carcinogenicity studies. None of the histopathologic changes are related to treatment with aspartame.

Prem Dua, D.V. M., Ph.D.

c.c.: Hattan (HFS-265)
Francke-Carroll (HFS-715)