Published in final edited form as:

*Toxicol Pathol*. 2010 August ; 38(5): 765–775. doi:10.1177/0192623310373777.

# Comparison of Historical Control Tumor Incidence Rates in Female Harlan Sprague-Dawley and Fischer 344/N Rats From Two-Year Bioassays Performed by the National Toxicology Program

**Gregg E. Dinse**[1], **Shyamal D. Peddada**[1], **Shawn F. Harris**[2], and **Susan A. Elmore**[3,4]

[1]Biostatistics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709 USA

[2]SRA International, Inc., 2605 Meridian Parkway, Durham, NC 27713 USA

[3]Cellular and Molecular Pathology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709 USA

[4]National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709 USA

## Abstract

The National Toxicology Program (NTP) has historically used Fischer 344/N (F344/N) rats for the majority of its bioassays. In 2008 the NTP switched rat strains and began using Harlan Sprague Dawley (SD) rats. The NTP had previously used female SD rats in nine bioassays. This article compares historical control (HC) tumor incidence rates from these nine SD rat studies with HC tumor rates from matched NTP F344/N rat bioassays to identify similarities and differences. Matching on sex, laboratory, diet, and route led to nine comparable F344/N rat studies. Our analyses revealed statistically significant strain differences in incidence rates for clitoral gland adenoma, mammary gland fibroadenoma, mammary gland carcinoma, thyroid gland C cell adenoma, and mononuclear cell leukemia. These represent five of the seven most common tumor types among female SD and F344/N rats in the NTP HC database. When vehicle was included as an additional matching criterion, the number of comparable F344/N rat studies dropped to four, but similar results were obtained. Among female F344/N rats, the incidence of pituitary gland pars distalis adenoma was significantly higher when the vehicle was corn oil as compared to water, suggesting a possible vehicle effect.

### Keywords

Fischer 344/N; historical control; multiple testing correction; NTP; Sprague Dawley; survival adjustment; tumor incidence

Address Correspondence to: Susan A. Elmore, NIEHS, P.O. Box 12233, Research Triangle Park, NC 27709, USA; elmore@niehs.nih.gov, 919-541-3474.

## Introduction

The inbred F344/N rat substrain has been used for the NTP rodent carcinogenic bioassay for more than 30 years. Over time, this specific colony of rats developed certain undesirable traits such as decreased fecundity, sporadic seizures and idiopathic chylothorax. Moreover, the spontaneous incidences of mononuclear cell leukemia and testicular interstitial cell tumors became unacceptably high. For example, according to the May 2009 NTP HC report (http://ntp.niehs.nih.gov; all routes/all vehicles), the average incidence rate for leukemia in males was 38.3% (536/1398; range 8%-58%) and in females was 21.3% (288/1350; range 8%-40%), and the average incidence rate for testicular interstitial cell tumors was 82.9% (1159/1398; range 58%-98%). Such high background tumor incidence rates and wide ranges in control animals may decrease the ability to detect subtle treatment-related effects. Thus, when a statistically significant increase in tumor incidence is found in test animals relative to concurrent controls, the apparent effect may not be considered biologically significant or treatment related if the tumor rate falls within its historical control range.

To address these problems with F344/N rats, the NTP researched other available rat strains and decided to use the outbred Wistar Han rat (King-Herbert and Thayer, 2006). This strain was chosen for its hybrid vigor, large litter size, long lifespan, resistance to disease, and low neonatal mortality. Many NTP studies call for the use of pregnant rats, with dosing commencing early in pregnancy. The NTP used Wistar Han rats for more than a year, and during that time it became evident that the reproductive robustness of the Wistar Han rat waned in a timed mating situation. Litter sizes were much smaller than the typical 10 pups/litter observed in natural mating situations. Because of smaller litter sizes and skewed sex ratios of the pups, the number of usable litters decreased significantly, negatively impacting the availability of animals for studies. The NTP subsequently considered purchasing Wistar Han rats from other venders. However, the large size (800 gram males) of the Wistar Han rats from other vendors precluded their use in NTP inhalation studies due to limits on cage size.

Consequently, the NTP decided to switch to the Harlan Sprague Dawley (Hsd:Sprague Dawley SD) rat for future bioassays (King-Herbert et al., 2009). Nine previous NTP bioassays had been performed using female Harlan SD rats to evaluate the chronic toxicity and carcinogenicity of dioxins and dioxin-like compounds. Historical data on spontaneous tumor rates in vehicle controls were available from those studies (http://ntp.niehs.nih.gov, retrieved 12/2009). The female SD rats previously used by the NTP were reasonably sized, between 375 and 400 grams peak weight, with an acceptable disease profile. They also produced large litters under timed (24 hour access) mating conditions with a good sex ratio, making them ideal for perinatal exposure studies. Additionally, this animal is commonly used in NTP reproductive and developmental studies, so its use as the “default” rat model would allow all metabolism and kinetic studies to be performed within the same strain.

The goal of this project was to compare the HC tumor incidence rates between F344/N and SD rats used in NTP bioassays, after accounting for potential sources of variability beyond strain. After matching on sex, laboratory, diet, route, and in some cases vehicle, formal

statistical analysis was used to compare these two strains with adjustments for body weight, survival, chronologic time, study-to-study variability, and multiple testing.

## Materials and Methods

### Studies

The data for our analyses were obtained from the control groups of 18 NTP chronic rodent carcinogenicity bioassays: nine SD rat bioassays and nine F344/N rat bioassays. We used all of the conventional NTP studies conducted in SD rats. Only females were used in seven of these nine SD rat studies; therefore we restricted our analyses to female rats. Female Harlan SD (Hsd:Sprague Dawley SD) rats were acquired from Harlan Sprague-Dawley, Inc., Indianapolis, IN. All nine SD rat studies were performed at Battelle Columbus Laboratory (Columbus, OH, USA), under contract to the NTP. In each study, the SD rats were fed the NTP-2000 diet, the route of exposure was gavage, and the vehicle was corn oil.

For comparison, we selected all F344/N rat studies that matched the SD rat studies with respect to sex (female), laboratory (Battelle Columbus), diet (NTP-2000) and route of exposure (gavage). This matching led to nine F344/N rat studies, where corn oil was the vehicle in four of these studies and water was the vehicle in the other five studies.

Fifty to fifty-three female rats were used in the control group of each core study (Table 1). This led to 473 SD rats and 450 F344/N rats, for a total of 923 rats. These rats were sacrificed when moribund or after two years on study. The NTP technical report numbers for these 18 studies are listed in Table 1, along with other summary information, including strain and vehicle. For access to the full technical reports and all of the data, see the NTP web site (http://ntp.niehs.nih.gov, retrieved 12/2009).

Animal handling and husbandry was conducted in accordance with NIH and IACUC policies and guidelines (http://www.iacuc.org/index.html, retrieved 12/31/2009). Animal studies were conducted in accordance with NTP two-year study protocol (http://ntp.niehs.nih.gov/go/9989, retrieved 12/31/2009) and NTP specifications (http://ntp.niehs.nih.gov/files/Specifications_2006Oct1.pdf).

### Data

Our primary focus was on strain differences in tumor incidence. We retrieved response information on many tumor classifications, though we excluded any with fewer than three occurrences in the entire NTP HC database because those data were considered too sparse. After excluding metastases and combinations of tumors, such as “carcinomas or adenomas” or “malignant tumors,” a total of 82 tumor types remained (Table 2).

In addition to tumor incidence data, we retrieved information on chronologic time, survival, and body weight. Chronologic time was summarized by the year in which each study began, and study-specific summaries of survival and body weight were calculated from individual animal data (Table 1). For each rat, time on study was recorded in days and body weight was measured in grams at various times during the experiment. Study-specific survival was summarized in terms of both the average life span and the proportion of rats surviving to the

end of the two-year bioassay. Study-specific body weight was summarized by the average body weight at one year on study. Only those body weights recorded between days 355 and 375 were averaged; if a rat was weighed multiple times during that period, the weight closest to day 365 was used; and any rat not weighed during that period, including those dying before one year, did not contribute to the average.

## Statistical Methods

We used a linear mixed model to analyze tumor incidence as a function of tumor type, rat strain, body weight, and chronologic time, while accounting for study-to-study variability. Rather than focusing directly on the empirical tumor rate, which is the number of tumor-bearing rats divided by the total number of rats, we incorporated the poly-3 survival adjustment of Bailer and Portier (1988). This survival adjustment, which accounts for the fact that not all rats had equal life spans, is important because rats that died early were at less risk of developing tumors than rats that died late. In addition, we applied an arcsine-root transformation to the survival-adjusted tumor rate to stabilize the variance of the response variable. This linear mixed model analysis was performed using Proc Mixed in the SAS software package (version 9.00, SAS Institute Inc., Cary, NC, USA).

Our analysis adjusted for the various explanatory factors as follows. We included a separate indicator variable for each of the 82 tumor types, each of the two rat strains, and each of the 18 studies. This approach is flexible and avoids assumptions about the exact form of the relationship between tumor response and the explanatory factors. Because some increases in tumor incidence rates are associated with increased body weight (Haseman et al., 2003), we wanted to adjust our strain comparisons for differences in body weight. Thus, for each control group, we incorporated a quantitative variable equal to the average body weight (in grams) after one year on study. To adjust for chronologic time, we also included a quantitative variable equal to the year the study started, which ranged from 1995 to 2004. We modeled the mean of the response variable as a linear function of these explanatory variables, plus strain-by-tumor interaction terms to allow different strain effects for different tumor types. Our analysis treated study, nested within strain, as a random effect to account for study-to-study variability. We treated all other factors as fixed effects.

Our analysis evaluated the statistical significance of each factor with an F-test from an analysis of variance (ANOVA). We also calculated a two-sided t-test based on least squares means to compare the strains for each tumor type. Each t-test assessed the null hypothesis that the two strains had equal mean responses, after adjusting for possible effects of study, chronologic time, and body weight. In a broad search for any tumor types having incidence rates that appeared to differ between female F344/N and SD rats, we examined the individual p-values for all 82 tumor types and considered any p-value below 0.05 as possibly being statistically significant. However, we also conducted a more conservative analysis, which incorporated a Bonferroni correction for multiple testing to account for having performed 82 tests. This stricter test declared a strain difference to be significant if the observed p-value was below 0.05/82 (or approximately 0.0006) rather than below the nominal 0.05 level, thus controlling the familywise error rate at 0.05.

We performed two sets of analyses. One analysis compared the nine SD rat studies with the nine F344/N rat studies that matched on all criteria except vehicle, and the other analysis compared the nine SD rat studies with the four F344/N rat studies that matched on all criteria, including vehicle (corn oil).

## Results

### Body Weight

The two strains of female rats differed with respect to body weight, with average study-specific weights at one year ranging from 255 to 293 grams for F344/N rats and from 317 to 345 grams for SD rats (Table 1). Among studies matched on sex, lab, diet, and route, the overall average body weight at one year was 276 grams for F344/N rats and 331 grams for SD rats (Table 3). When further matched on corn oil vehicle, the average one-year body weights did not change appreciably.

### Survival

The two strains also exhibited different mortality patterns, with female F344/N rats tending to live longer than female SD rats. Among the 18 NTP bioassays matched on sex, lab, diet, and route, a greater proportion of F344/N rats (range 60%-76%) per control group survived to the end of the two-year study than did their SD counterparts (range 28%-51%; Table 1), with respective overall means of 67% and 43% (Table 3). Similarly, the average time on study ranged from 668 to 710 days for the F344/N rats, compared with 567 to 656 days for the SD rats (Table 1), with respective overall means of 688 and 630 days (Table 3). Among the subset of four F344/N rat studies with the same vehicle (corn oil) as the nine SD rat studies, the proportion surviving two years was 62%-68% and the average lifespan was 668-699 days. The observed variation in survival, both within and between strains, illustrates the need for a survival adjustment in the statistical analysis of the tumor rates.

### Tumor Incidence

Several tumor types were common in both strains of female rats (Table 2). Among all 18 studies, there were four tumor types for which both strains had incidence rates above 10%: mammary gland fibroadenoma (48% in F344/N and 67% in SD), pituitary gland pars distalis adenoma (45% in F344/N and 39% in SD), thyroid gland C cell adenoma (14% in F344/N and 25% in SD), and uterine stromal polyp (14% in F344/N and 14% in SD). There were also two tumor types with incidence rates of 10% or higher in one strain but not the other: mammary gland carcinoma (2% in F344/N and 10% in SD) and mononuclear cell leukemia (17% in F344/N and 1% in SD). The incidence rates for the remaining 76 tumor types were below 10% in both strains. In fact, the rates were only 1% or less in both strains for 70 of the 82 tumor types considered. Each tabulated percentage is the observed number of tumor-bearing rats divided by the total number of rats examined for that tumor type.

### Explanatory Factors

The only factors that had a statistically significant impact on tumor incidence were tumor type and the interaction between strain and tumor type (each with $p < 0.0001$). All other factors had p-values above 0.15 and thus were not considered statistically significant.

## Strain Differences

As a first step toward identifying which tumor types exhibited different incidence rates between female F344/N and SD rats, we calculated a p-value for each of the 82 tumor types, based on all 18 NTP bioassays that matched on sex, lab, diet and route (Table 2). Initially, we viewed any p-value below 0.05 as statistically significant and worthy of further attention. This criterion flagged 14 tumor types as possibly having different strain-specific incidence rates (see Table 2 and the first column of Table 4). The p-values below 0.05 are indicated with an asterisk.

Several tumor types had identical strain-specific incidence rates, but the p-values for comparing strains differed. For example, the incidence of malignant lymphoma was 1/450 in F344/N rats and 6/473 in SD rats, as was the incidence of uterine carcinoma, but the corresponding p-values were 0.139 and 0.037, respectively. This phenomenon occurred because the rates for the two tumor types were based on animals from different studies, and the p-values were obtained from an analysis that adjusted for factors (e.g., body weight and calendar time) that varied across studies.

When we also matched on vehicle and compared the nine SD rat studies with the four F344/N rat studies that used corn oil as the vehicle, there were 10 tumor types with a p-value below 0.05 (Table 4, second column). Of these 10 tumor types, nine were on the previous list based on all 18 studies and one was not. Those on both lists include: clitoral gland adenoma, clitoral gland carcinoma, lung alveolar/bronchiolar adenoma, mammary gland carcinoma, mammary gland fibroadenoma, mononuclear cell leukemia, malignant mesothelioma, thyroid gland C cell adenoma, and thyroid gland C cell carcinoma. The only strain difference that was newly identified after matching on vehicle was for pituitary gland pars distalis adenoma. There were five tumor types, however, with p-values that rose above 0.05 (and thus provided less evidence of a strain difference) after matching on vehicle: adrenal medulla pheochromocytoma benign, pituitary gland pars intermedia adenoma, skin basal cell adenoma, thyroid gland follicular cell carcinoma, and uterine carcinoma.

At the next stage, to guard against false positives, we incorporated a Bonferroni correction for multiple testing to account for having performed 82 tests. This correction produced a much stricter test, which only declared a strain difference to be significant if the observed p-value was below 0.05/82 (or approximately 0.0006) rather than below the nominal 0.05 level. The corrected test identified strain differences in incidence for seven tumor types when using all 18 studies (which matched on sex, lab, diet and route) and six tumor types based on the subset of 13 studies that also matched on vehicle. The p-values below the Bonferroni cutoff are indicated with a “B” in Tables 2 and 4. Five tumor types were significant whether matching on vehicle or not: clitoral gland adenoma, mammary gland carcinoma, mammary gland fibroadenoma, mononuclear cell leukemia, and thyroid gland C cell adenoma. Two tumor types lost their significance after matching on vehicle: adrenal medulla pheochromocytoma benign and thyroid gland C cell carcinoma. One tumor type became significant only after matching on vehicle: lung alveolar/bronchiolar adenoma (Table 4).

### Vehicle Effects in F344/N Rats

For most tumor types, the statistical significance of the strain difference in incidence did not depend on whether or not we matched on vehicle. That is, whether the nine corn oil SD rat studies were compared with only the four corn oil F344/N rat studies or with both the four corn oil F344/N rat studies and the five water F344/N rat studies, the results were usually similar. In a few cases, however, the vehicle used in the F344/N rat studies appeared to have an effect and we decided to investigate further. Thus, for each tumor type, we used a two-sided t-test to assess the equality of the responses in the five water and four corn oil F344/N rat studies (Table 4, third column). Of the 15 tumor types flagged by one of our previous analyses, four had an unadjusted p-value for this vehicle comparison that was below 0.05, though only one was still significant after adjusting for multiple testing. The one tumor type that appeared to depend on which vehicle was used in female F344/N rats was pituitary gland pars distalis adenoma. The three other tumor types that were suggestive of a vehicle effect were: adrenal medulla pheochromocytoma benign, lung alveolar/bronchiolar adenoma, and mononuclear cell leukemia.

## Discussion

Since the NTP was established in 1978, the F344/N rat has been the default strain of rat for the majority of its carcinogenicity bioassays. During the past 30 years, the NTP has maintained a historical control database with strain- and sex-specific spontaneous tumor incidence data from control animals. This database is updated annually to consist of data from all control animals from two-year bioassays within the most recent five-year window. Although the concurrent control tumor incidence data is always the most appropriate to use when evaluating the significance of tumor rates among treated groups, there are certain situations in which the use of historical control data can be helpful, such as the interpretation of rare tumors and marginally increased tumor incidences (Deschl et al., 2002; Greim et al., 2003; Keenan et al., 2009a).

When the NTP switched from the F344/N rat to the SD rat for the majority of its carcinogenicity bioassays, there was a question of how the control tumor incidence data might differ between these two strains. The NTP had previously used female SD rats for the evaluation of nine bioassays that involved dioxins and dioxin-like compounds. These were primarily mechanistic studies and, for this group of chemicals, the SD rat was chosen for two reasons: previous studies of dioxins and dioxin-like compounds had used SD rats and choosing a strain with a lower incidence rate of mononuclear cell leukemia would simplify the interpretation of complex hepatic lesions (Brix et al., 2005; Walker et al., 2005; Hailey et al., 2005). Control tumor incidence data were collected from these nine SD rat studies and were therefore available within the NTP HC database for comparison to matched F344/N rat data.

### Challenges in Comparing Tumor Rates

When comparing spontaneous tumor rates, especially from different sources, it is crucial to match on as many experimental conditions as possible and to adjust for suspected risk factors to avoid bias (Haseman, 1995). In addition, rather than simply controlling the usual

Type I error rate (or false positive rate), we want to control the familywise error rate (FWER) when testing multiple hypotheses (Hochberg and Tamhane, 1987). The FWER is the probability of rejecting at least one true null hypothesis among all hypotheses tested. For example, suppose we are interested in testing 25 independent hypotheses, each at a Type I error rate of 0.05. The probability of rejecting at least one exceeds 0.72. Thus, there is roughly a 72% chance of falsely rejecting at least one true null hypothesis by chance alone. Ideally we would like to limit the FWER to a small probability, such as 0.05. To accomplish this, we use the Bonferroni method, which controls the FWER at the desired level, even if the tests (or tumors) are not independent.

In addressing these many challenges, our comparison of F344/N and SD rat tumor incidence data controlled for potential sources of variation by matching on sex, lab, diet, route, and in some cases vehicle, and by adjusting for body weight, survival, chronologic time, and study-to-study variability. Also, we made a conservative (Bonferroni) correction for multiple testing by performing each of the 82 comparisons at the 0.05/82 significance level to keep the familywise error rate at or below 0.05.

## Strain Differences

Eight tumors had incidence rates greater than 5% in either the F344/N or SD rats: adrenal medulla benign pheochromocytoma, clitoral gland adenoma, mammary gland carcinoma, mammary gland fibroadenoma, mononuclear cell leukemia, pituitary gland pars distalis adenoma, thyroid gland C cell adenoma, and uterus polyp (Table 2). Of these eight most common tumors, five had statistically significant strain differences in incidence rates after matching on all criteria (sex, lab, diet, route, vehicle) and adjusting for multiple testing: clitoral gland adenoma, mammary gland carcinoma, mammary gland fibroadenoma, mononuclear cell leukemia, and thyroid gland C cell adenoma (Table 5). After relaxing the stringency of our comparison by not matching on vehicle, the difference in incidence rates for adrenal medulla benign pheochromocytoma became significant (Table 5). Alternatively, if we matched on all criteria but did not adjust for multiple testing, the difference in incidence rates for pituitary gland pars distalis adenoma became significant (Table 5). Uterine polyps were common in both F344/N (14.4%) and SD (13.7%) rats, but the strain differences were not statistically significant, even without adjusting for multiple testing, whether we matched on all criteria or all except vehicle (Table 2).

Several of the less common tumors also suggested strain differences. After adjusting for multiple testing, strain differences for lung alveolar/bronchiolar adenoma were significant when matching on all criteria, despite not being significant when matching on all criteria except vehicle (Table 5). Without a correction for multiple testing, these differences were significant whether matching on sex, lab, diet, and route only, or also including vehicle. In contrast, the strain differences for thyroid gland C cell carcinoma were significant with or without the Bonferroni correction when matching on all criteria except vehicle, but only without the multiple testing adjustment when including vehicle (Table 5).

None of the other tumors showed significant evidence of strain differences after adjusting for multiple testing, though some suggested strain differences if the Bonferroni correction was not performed. In this latter category, clitoral gland carcinoma and malignant

mesothelioma suggested strain differences, whether matching on all criteria or all criteria except vehicle, whereas pituitary gland pars intermedia adenoma, skin basal cell adenoma, thyroid gland follicular cell carcinoma, and uterus carcinoma only suggested a strain difference when the matching criteria did not include vehicle (Table 5).

## Corn Oil Vehicle Effect on Pituitary Gland Pars Distalis Adenoma

In the absence of a correction for multiple testing, the difference in incidence rates for pituitary gland pars distalis adenoma between female F344/N (201/450, 45%) and SD (183/471, 39%) rats was significant after matching on all criteria, including vehicle, but not if vehicle was dropped as a matching criterion (Table 5), suggesting a possible vehicle effect in female F344/N rats. Among female F344/N rats, the incidence rates by vehicle were 106/200 (53%) for corn oil and 95/250 (38%) for water, which are significantly different, even after applying a Bonferroni correction for multiple testing (Table 4, last column).

Neoplasms of the pars distalis are one of the most common neoplasms in the laboratory rat (MacKenzie and Boorman, 1990). The cause of spontaneous neoplasms of the pars distalis is unknown, but hormonal imbalances related to aging or stress may be a factor (Greim et al., 2003; Attia, 1985). Estrogen has a trophic effect on the development of spontaneous neoplasms of the pars distalis and the F344 rat is known to be more sensitive to estrogen-induced hyperplasia of the pars distalis than the Sprague Dawley rat (Fujimoto et al., 1987; MacKenzie and Boorman, 1990). The incidence of pituitary neoplasms in F344/N rats is also reported to be directly correlated with body weight (Haseman et al., 1997; Gries and Young, 1982). Dietary restriction results in significant reductions in body weight and is also reported to reduce the incidence of spontaneous pituitary neoplasms in laboratory rats (Keenan et al, 1994).

Haseman et al. performed a retrospective study of NTP carcinogenicity bioassays in 1985 to determine if F344/N rats receiving corn oil by gavage showed tumor incidences that differed from those of untreated control animals (Haseman et al., 1985). The reported effects of corn oil gavage on rats were increases in body weight and survival and pancreatic acinar cell tumors and decreases in mononuclear cell leukemia, all occurring in male but not female F344 rats. Some of these effects in male rats appear to be inter-related. Since corn oil gavage reduces leukemia, the male rats live longer. Similarly, the increased rates of pancreatic acinar cell tumors may be a direct effect of corn oil, in combination with other dietary fatty acids, on the metabolism of initiated pancreatic cells (Haseman and Rao, 1992).

Previous investigations have not shown vehicle to affect the incidence rates of pituitary neoplasms in female F344/N rats from NTP studies (Haseman and Rao 1992; Haseman et al. 1985). However, there is evidence that decreased body weight can reduce the incidence of pituitary gland tumors in both male and female F344/N rats (Haseman et al., 1997; Haseman et al., 2003). In the current evaluation, the female F344/N rats were fed the NTP-2000 diet and gavaged with either corn oil or water, but had similar mean body weights at one year on study: 278.9g for corn oil gavage studies and 274.5g for water gavage studies. Although the incidence of pituitary gland tumors in male and female F344/N rats has been directly correlated with body weight (Haseman et al., 1997; Haseman et al., 2003), in this study corn oil was not associated with an increase in body weight. The lack of such a correlation in our

analysis suggests that the increase in pituitary gland pars distalis adenomas in female F344/N rats may have been a direct effect of the corn oil.

## Pooling Across Vehicles

In some specific cases involving rare tumors, pooling incidence rates across vehicles may provide greater power to detect differences between strains. For example, the incidence of uterus carcinoma was 6/473 (1.27%) in SD rats given corn oil gavage, 1/200 (0.50%) in F344/N rats given corn oil gavage, and 0/250 (0.00%) in F344/N rats given water gavage. By matching on all criteria, including vehicle, we found no significant difference in tumor rates between SD and F344/N rats ($p = 0.248$; Table 4). However, when we did not match on vehicle, the strain difference became significant ($p = 0.037$; Table 4). This change in significance was due to the sample size for the F344/N group essentially increasing from 200 to 450, resulting in the statistical test having greater power.

## Comparison to Previous Analysis

In 2005 Brix et al. reported spontaneous tumor incidence data in female control SD rats and also compared tumor incidence data for six common tumors between SD and F344/N rats used in NTP studies. They informally noted similarities and differences among those six tumor types, but did not perform formal statistical comparisons of the two strains. Some of their incidence rates agree well with ours and others do not, but this lack of consistency was expected for several reasons. Brix et al. (2005) combined adenomas and carcinomas in their comparisons of clitoral gland neoplasms, pituitary gland pars distalis neoplasms, and thyroid gland C cell neoplasms, whereas we listed adenomas and carcinomas separately. They used data on 371 SD rats from seven NTP studies, whereas we used data on 473 SD rats from those same studies plus two additional NTP studies. Among studies using the NTP-2000 diet, they compared SD rats from gavage studies with F344/N rats from feed studies, whereas we held route constant by using only gavage studies for both SD and F344/N rats. Brix et al. (2005) also presented incidence rates from gavage studies in F344/N rats, but those rats were fed the NIH-07 diet, whereas the SD rats were fed the NTP-2000 diet. Thus, our formal comparisons matched on sex, lab, diet and route (and in some cases vehicle), whereas their informal comparisons either matched on sex and diet (but not lab, route and vehicle) or else on sex, route and vehicle (but not lab and diet).

## Concluding Remarks

The major concern with using historical control data from different studies or sources is the comparability of the study under evaluation with the studies in the historical control database, in light of known and unknown sources of variability (Keenan et al., 2009a; Keenan et al., 2009b; Haseman and Rao, 1992). Potential sources of variability in chronic rodent bioassays include, but are not limited to: species, strain, age, sex, laboratory, dietary factors, route of exposure, vehicle, body weight, survival, animal room environment, gross necropsy, slide preparation techniques, and histopathology diagnoses (Haseman et al., 1989; Rao et al., 1987; Rao and Crockett, 2003; Haseman and Rao 1992). As shown in this investigation, sources of variability within a strain, such as diet, vehicle and body weight, can be inter-related and can affect specific tumor rates, such as pituitary gland tumors. Furthermore, as we illustrated, pooling of tumor data across vehicles within a strain may be

reasonable for rare tumors, such as uterus carcinoma. However, whether tumors are rare or common, it is not appropriate to combine data across strains because the strain effect may be significant.

## Acknowledgments

This research was supported [in part] by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (Z01ES101744; Z01ES045007). We are grateful for the constructive comments from Michelle Hooth, Grace Kissling, and David Malarkey.

## References

Attia MA. Neoplastic and nonneoplastic lesions in aging female rats with special reference to the functional morphology of the hyperplastic and neoplastic changes in the pituitary gland. Arch Toxicol. 1985; 57:77–83. [PubMed: 2992416]

Bailer AJ, Portier CJ. Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples. Biometrics. 1988; 44:417–31. [PubMed: 3390507]

Brix AE, Nyska A, Haseman JK, Sells DM, Jokinen MP, Walker NJ. Incidences of selected lesions in control female Harlan Sprague-Dawley rats from two-year studies performed by the National Toxicology Program. Toxicol Pathol. 2005; 33:477–83. [PubMed: 16036865]

Deschl U, Kittel B, Rittinghausen S, Morawietz G, Kohler M, et al. The value of historical control data-scientific advantages for pathologists, industry and agencies. Toxicol Pathol. 2002; 30:80–7. [PubMed: 11890480]

Fujimoto M, Yoshino E, Hirakawa K, Chihara K, Ibata Y. Studies on estrogen induced pituitary tumor in the rat with special reference to the relationship of the tuberoinfundibular dopamine neuron system. J Neurooncol. 1987; 5:151–9. [PubMed: 3668610]

Greim H, Gelbke HP, Reuter U, Thielmann HW, Edler L. Evaluation of historical control data in carcinogenicity studies. Hum Exp Toxicol. 2003; 22:541–9. [PubMed: 14655720]

Gries CL, Young SS. Positive correlation of body weight with pituitary tumor incidence in rats. Fundam Appl Toxicol. 1982; 2:145–8. [PubMed: 7185611]

Hailey JR, Walker NJ, Sells DM, Brix AE, Jokinen MP, Nyska A. Classification of proliferative hepatocellular lesions in harlan sprague-dawley rats chronically exposed to dioxin-like compounds. Toxicol Pathol. 2005; 33:165–74. [PubMed: 15805068]

Haseman JK. Data analysis: statistical analysis and use of historical control data. Regul Toxicol Pharmacol. 1995; 21:52–9. discussion 81-6. [PubMed: 7784636]

Haseman JK, Huff JE, Rao GN, Arnold JE, Boorman GA, McConnell EE. Neoplasms observed in untreated and corn oil gavage control groups of F344/N rats and (C57BL/6N X C3H/HeN)F1 (B6C3F1) mice. J Natl Cancer Inst. 1985; 75(5):975–84. [PubMed: 3863995]

Haseman JK, Huff JE, Rao GN, Eustis SL. Sources of variability in rodent carcinogenicity studies. Fundam Appl Toxicol. 1989; 12:793–804. [PubMed: 2744280]

Haseman JK, Ney E, Nyska A, Rao GN. Effect of diet and animal care/housing protocols on body weight, survival, tumor incidences, and nephropathy severity of F344 rats in chronic studies. Toxicol Pathol. 2003; 31:674–81. [PubMed: 14585736]

Haseman JK, Rao GN. Effects of corn oil, time-related changes, and inter-laboratory variability on tumor occurrence in control Fischer 344 (F344/N) rats. Toxicol Pathol. 1992; 20:52–60. [PubMed: 1411131]

Haseman JK, Young E, Eustis SL, Hailey JR. Body weight-tumor incidence correlations in long-term rodent carcinogenicity studies. Toxicol Pathol. 1997; 25:256–263. [PubMed: 9210256]

Hochberg, Y.; Tamhane, AC. Multiple Comparison Procedures. John Wiley and Sons; New York, NY: 1987.

Keenan C, Elmore S, Francke-Carroll S, Kemp R, Kerlin R, Peddada S, Pletcher J, Rinke M, Schmidt SP, Taylor I, Wolf DC. Best practices for use of historical control data of proliferative rodent lesions. Toxicol Pathol. 2009a; 37:679–93. [PubMed: 19454599]

Keenan C, Elmore S, Francke-Carroll S, Kerlin R, Peddada S, Pletcher J, Rinke M, Schmidt SP, Taylor I, Wolf DC. Potential for a global historical control database for proliferative rodent lesions. Toxicol Pathol. 2009b; 37:677–8. [PubMed: 19638441]

Keenan KP, Smith PF, Hertzog P, Soper K, Ballam GC, Clark RL. The effects of overfeeding and dietary restriction on Sprague-Dawley rat survival and early pathology biomarkers of aging. Toxicol Pathol. 1994; 22:300–15. [PubMed: 7817120]

King-Herbert A, Sills RC, Bucher JR. Commentary: update on animal models for NTP studies. Toxicol Pathol. 2009; 000:1–2.

King-Herbert A, Thayer K. NTP workshop: animal models for the NTP rodent cancer bioassay: stocks and strains--should we switch? Toxicol Pathol. 2006; 34:802–5. [PubMed: 17162538]

MacKenzie, WF.; Boorman, GA. Pituitary gland. In: Boorman, GA.; Eustis, SL.; Elwell, MR.; Montgomery, CA.; MacKenzie, WF., editors. Pathology of the Fischer Rat Reference and Atlas. Academic Press, Inc.; San Diego, CA: 1990. p. 485-500.

Rao GN, Crockett PW. Effect of diet and housing on growth, body weight, survival and tumor incidences of B6C3F1 mice in chronic studies. Toxicol Pathol. 2003; 31:243–50. [PubMed: 12696586]

Rao GN, Piegorsch WW, Haseman JK. Influence of body weight on the incidence of spontaneous tumors in rats and mice of long-term studies. Am J Clin Nutr. 1987; 45:252–60. [PubMed: 3799516]

Walker NJ, Crockett PW, Nyska A, Brix AE, Jokinen MP, Sells DM, Hailey JR, Easterling M, Haseman JK, Yin M, Wyde ME, Bucher JR, Portier CJ. Dose-additive carcinogenicity of a defined mixture of “dioxin-like compounds”. Environ Health Perspect. 2005; 113:43–8. [PubMed: 15626646]

## Abbreviations

| | |
|---|---|
| **NTP** | National Toxicology Program |
| **F344/N** | Fischer 344/N |
| **SD** | Sprague Dawley |
| **HC** | historical control |
| **FWER** | familywise error rate |

**Table 1**

Summary information for control female rats from 18 long-term carcinogenicity bioassays conducted by the U.S. National Toxicology Program (NTP), matched on laboratory (Battelle Columbus), diet (NTP-2000), and route (gavage).

| Strain | Vehicle | Technical Report Number | Year Study Started | Number of Rats | Mean 1-year Body Weight (grams) | Percent (%) Surviving to End of Study | Mean Life Span (days) |
|---|---|---|---|---|---|---|---|
| Sprague Dawley | corn oil | 520 | 1998 | 53 | 327 | 28 | 567 |
| Sprague Dawley | corn oil | 521 | 1998 | 53 | 329 | 47 | 637 |
| Sprague Dawley | corn oil | 525 | 1999 | 53 | 317 | 47 | 655 |
| Sprague Dawley | corn oil | 526 | 1998 | 53 | 326 | 30 | 616 |
| Sprague Dawley | corn oil | 529 | 1998 | 53 | 341 | 45 | 641 |
| Sprague Dawley | corn oil | 530 | 1998 | 53 | 341 | 42 | 629 |
| Sprague Dawley | corn oil | 531 | 1999 | 53 | 345 | 51 | 656 |
| Sprague Dawley | corn oil | 558 | 2003 | 50 | 323 | 50 | 652 |
| Sprague Dawley | corn oil | 559 | 2004 | 52 | 329 | 40 | 616 |
| Fischer 344/N | corn oil | 551 | 2002 | 50 | 278 | 66 | 699 |
| Fischer 344/N | corn oil | 557 | 2002 | 50 | 270 | 62 | 668 |
| Fischer 344/N | corn oil | 563 | 2003 | 50 | 281 | 62 | 682 |
| Fischer 344/N | corn oil | 571 | 2004 | 50 | 286 | 68 | 683 |
| Fischer 344/N | water | 497 | 1995 | 50 | 255 | 76 | 710 |
| Fischer 344/N | water | 512 | 1997 | 50 | 259 | 60 | 672 |
| Fischer 344/N | water | 541 | 2001 | 50 | 293 | 76 | 703 |
| Fischer 344/N | water | 554 | 2002 | 50 | 283 | 62 | 679 |
| Fischer 344/N | water | 567 | 2004 | 50 | 284 | 72 | 700 |

## Table 2

Tumor rates (percentages) and p-values for strain differences between control female Fischer 344/N and Hsd: Sprague Dawley SD rats in 18 NTP studies matched on laboratory (Battelle Columbus), diet (NTP-2000), and route (gavage).

| | Tumor Rate (%) | | |
|---|---|---|---|
| **Organ/Morphology** | **Fischer 344/N** | **Sprague Dawley** | **p-value[a]** |
| Adrenal Cortex, Adenoma | 3/450 (0.67) | 5/471 (1.06) | 0.645 |
| Adrenal Cortex, Carcinoma | 0/450 (0.00) | 3/471 (0.64) | 0.217 |
| Adrenal Medulla, Pheochromocytoma, Benign | 13/450 (2.89) | 34/470 (7.23) | < 0.001*[B] |
| Adrenal Medulla, Pheochromocytoma, Complex | 0/450 (0.00) | 1/470 (0.21) | 0.926 |
| Adrenal Medulla, Pheochromocytoma, Malignant | 1/450 (0.22) | 1/470 (0.21) | 0.662 |
| Bone, Osteosarcoma | 1/450 (0.22) | 1/473 (0.21) | 0.718 |
| Brain, Astrocytoma, Malignant | 0/450 (0.00) | 2/473 (0.42) | 0.661 |
| Brain, Glioma, Malignant | 0/450 (0.00) | 0/473 (0.00) | 0.650 |
| Brain, Oligodendroglioma, Malignant | 1/450 (0.22) | 1/473 (0.21) | 0.720 |
| Clitoral Gland, Adenoma | 26/447 (5.82) | 1/465 (0.22) | < 0.001*[B] |
| Clitoral Gland, Carcinoma | 6/447 (1.34) | 0/465 (0.00) | 0.003* |
| Ear, Neural Crest Tumor | 1/450 (0.22) | 1/473 (0.21) | 0.670 |
| Forestomach, Squamous Cell Carcinoma | 1/450 (0.22) | 2/473 (0.42) | 0.911 |
| Forestomach, Squamous Cell Papilloma | 1/450 (0.22) | 0/473 (0.00) | 0.320 |
| Heart, Schwannoma, Benign | 1/450 (0.22) | 0/471 (0.00) | 0.320 |
| Heart, Schwannoma, Malignant | 1/450 (0.22) | 3/471 (0.64) | 0.704 |
| Intestine Large, Colon/Rectum, Adenoma | 1/450 (0.22) | 0/473 (0.00) | 0.320 |
| Intestine Small, Jejunum, Leiomyosarcoma | 0/450 (0.00) | 2/473 (0.42) | 0.480 |
| Kidney, Renal Tubule, Adenoma | 0/450 (0.00) | 2/472 (0.42) | 0.506 |
| Kidney, Renal Tubule, Carcinoma | 0/450 (0.00) | 1/472 (0.21) | 0.894 |
| Kidney, Renal Tubule, Lipoma | 0/450 (0.00) | 2/472 (0.42) | 0.509 |
| Kidney, Renal Tubule, Nephroblastoma | 1/450 (0.22) | 2/472 (0.42) | 0.835 |
| Kidney, Renal Tubule, Sarcoma | 0/450 (0.00) | 0/472 (0.00) | 0.650 |
| Liver, Hepatocellular, Adenoma | 4/450 (0.89) | 6/473 (1.27) | 0.610 |
| Lung, Alveolar/Bronchiolar, Adenoma | 9/450 (2.00) | 2/471 (0.42) | 0.007* |
| Lung, Alveolar/Bronchiolar, Carcinoma | 1/450 (0.22) | 0/471 (0.00) | 0.330 |
| Lung, Squamous Cell Carcinoma | 1/450 (0.22) | 0/471 (0.00) | 0.318 |
| Mammary Gland, Adenoma | 9/450 (2.00) | 12/473 (2.54) | 0.234 |
| Mammary Gland, Carcinoma | 11/450 (2.44) | 48/473 (10.15) | < 0.001*[B] |
| Mammary Gland, Fibroadenoma | 218/450 (48.44) | 319/473 (67.44) | < 0.001*[B] |
| Mesentery, Schwannoma, Malignant | 0/450 (0.00) | 1/473 (0.21) | 0.919 |
| Multiple Organs, Histiocytic Sarcoma | 0/450 (0.00) | 2/473 (0.42) | 0.478 |
| Multiple Organs, Leukemia, Mononuclear Cell | 75/450 (16.67) | 4/473 (0.85) | < 0.001*[B] |
| Multiple Organs, Lymphoma, Malignant[b] | 1/450 (0.22) | 6/473 (1.27) | 0.139 |
| Multiple Organs, Mesothelioma, Malignant | 4/450 (0.89) | 0/473 (0.00) | 0.023* |

| | Tumor Rate (%) | | |
|---|---|---|---|
| **Organ/Morphology** | **Fischer 344/N** | **Sprague Dawley** | **p-value[a]** |
| Oral Mucosa, Squamous Cell Carcinoma | 2/450 (0.44) | 4/473 (0.85) | 0.476 |
| Oral Mucosa, Squamous Cell Papilloma | 0/450 (0.00) | 0/473 (0.00) | 0.650 |
| Ovary, Cystadenoma | 1/449 (0.22) | 1/469 (0.21) | 0.674 |
| Ovary, Granulosa Cell Tumor, Benign | 1/449 (0.22) | 0/469 (0.00) | 0.315 |
| Ovary, Granulosa Cell Tumor, Malignant | 1/449 (0.22) | 3/469 (0.64) | 0.475 |
| Ovary, Granulosa-Theca Tumor, Malignant | 0/449 (0.00) | 0/469 (0.00) | 0.650 |
| Ovary, Luteoma | 0/449 (0.00) | 3/469 (0.64) | 0.216 |
| Ovary, Tubulostromal, Adenoma | 0/449 (0.00) | 0/469 (0.00) | 0.650 |
| Pancreas, Acinar Cell, Adenoma | 1/450 (0.22) | 1/468 (0.21) | 0.713 |
| Pancreas, Islet Cell, Adenoma | 5/450 (1.11) | 5/469 (1.07) | 0.777 |
| Pancreas, Islet Cell, Carcinoma | 0/450 (0.00) | 3/469 (0.64) | 0.212 |
| Parathyroid Gland, Adenoma | 4/420 (0.95) | 1/429 (0.23) | 0.079 |
| Pituitary Gland, Pars Distalis, Adenoma | 201/450 (44.67) | 183/471 (38.85) | 0.996 |
| Pituitary Gland, Pars Distalis, Carcinoma | 5/450 (1.11) | 3/471 (0.64) | 0.170 |
| Pituitary Gland, Pars Intermedia, Adenoma | 0/450 (0.00) | 5/471 (1.06) | 0.037* |
| Skeletal Muscle, Rhabdomyosarcoma | 0/450 (0.00) | 1/473 (0.21) | 0.914 |
| Skin, Basal Cell, Adenoma | 5/450 (1.11) | 1/473 (0.21) | 0.010* |
| Skin, Basal Cell, Carcinoma | 0/450 (0.00) | 1/473 (0.21) | 0.915 |
| Skin, Fibroma | 9/450 (2.00) | 8/473 (1.69) | 0.734 |
| Skin, Fibrosarcoma | 1/450 (0.22) | 1/473 (0.21) | 0.676 |
| Skin, Keratoacanthoma | 4/450 (0.89) | 2/473 (0.42) | 0.474 |
| Skin, Neural Crest Tumor | 0/450 (0.00) | 0/473 (0.00) | 0.650 |
| Skin, Sarcoma | 0/450 (0.00) | 0/473 (0.00) | 0.650 |
| Skin, Schwannoma, Malignant | 1/450 (0.22) | 2/473 (0.42) | 0.869 |
| Skin, Squamous Cell Carcinoma | 0/450 (0.00) | 0/473 (0.00) | 0.650 |
| Skin, Squamous Cell Papilloma | 3/450 (0.67) | 1/473 (0.21) | 0.136 |
| Skin, Trichoepithelioma | 1/450 (0.22) | 0/473 (0.00) | 0.329 |
| Spleen Hemangiosarcoma | 2/450 (0.44) | 0/469 (0.00) | 0.224 |
| Thymus, Thymoma, Benign | 1/433 (0.23) | 1/461 (0.22) | 0.646 |
| Thyroid Gland, C-Cell Adenoma | 61/450 (13.56) | 119/468 (25.43) | < 0.001*[B] |
| Thyroid Gland, C-Cell Carcinoma | 6/450 (1.33) | 18/468 (3.85) | < 0.001*[B] |
| Thyroid Gland, Follicular Cell Adenoma | 2/450 (0.44) | 2/468 (0.43) | 0.719 |
| Thyroid Gland, Follicular Cell Carcinoma | 5/450 (1.11) | 1/468 (0.21) | 0.024* |
| Tongue, Squamous Cell Carcinoma | 0/450 (0.00) | 0/473 (0.00) | 0.650 |
| Tongue, Squamous Cell Papilloma | 1/450 (0.22) | 0/473 (0.00) | 0.315 |
| Tooth, Odontoma | 1/450 (0.22) | 1/473 (0.21) | 0.687 |
| Urinary Bladder, Papilloma | 0/450 (0.00) | 1/469 (0.21) | 0.908 |
| Uterus, Adenoma | 0/450 (0.00) | 2/473 (0.42) | 0.748 |
| Uterus, Carcinoma | 1/450 (0.22) | 6/473 (1.27) | 0.037* |
| Uterus, Hemangioma | 0/450 (0.00) | 0/473 (0.00) | 0.650 |

| | Tumor Rate (%) | | |
|---|---|---|---|
| **Organ/Morphology** | **Fischer 344/N** | **Sprague Dawley** | **p-value[a]** |
| Uterus, Leiomyosarcoma | 1/450 (0.22) | 1/473 (0.21) | 0.684 |
| Uterus, Polyp, Stromal | 65/450 (14.44) | 65/473 (13.74) | 0.829 |
| Uterus, Sarcoma, Stromal | 1/450 (0.22) | 0/473 (0.00) | 0.323 |
| Uterus, Schwannoma, Malignant | 1/450 (0.22) | 4/473 (0.85) | 0.190 |
| Vagina, Polyp | 1/450 (0.22) | 1/473 (0.21) | 0.704 |
| Vagina, Sarcoma | 0/450 (0.00) | 1/473 (0.21) | 0.862 |
| Zymbal's Gland, Carcinoma | 1/450 (0.22) | 1/473 (0.21) | 0.668 |

[a] An asterisk (*) indicates the p-value for a strain difference is statistically significant at the 0.05 level, in the absence of a correction for multiple testing. An additional superscript B indicates the p-value for a strain difference remains statistically significant after applying a Bonferroni correction for multiple testing, to adjust for 82 tests having been performed.

[b] Malignant lymphoma includes histiocytic, lymphocytic, mixed, not otherwise specified, and undifferentiated cell types.

## Table 3

Comparison of Fischer 344/N and Hsd: Sprague Dawley SD rats with respect to mean body weight, survival percentage, and life span in 18 NTP studies (nine per strain), matched on laboratory (Battelle Columbus), diet (NTP-2000), and route (gavage).

| | Mean ± Standard Deviation for: | | |
|---|---|---|---|
| **Strain** | **1-Year Body Weight (grams)** | **Percent Surviving to End of Study** | **Life Span (days)** |
| Fischer 344/N | 276.4 ± 12.0 | 67.2 ± 5.9 | 688.4 ± 14.1 |
| Sprague Dawley | 331.1 ± 8.8 | 42.5 ± 7.7 | 629.9 ± 26.5 |

## Table 4

Strain and vehicle comparisons with respect to the incidence of select tumor types among control female Fischer 344/N and Hsd: Sprague Dawley SD rats in NTP studies matched on laboratory (Battelle Columbus), diet (NTP-2000), and route (gavage).

| | P-values[a] for the indicated comparisons (number of studies): | | |
|---|---|---|---|
| | SD, corn oil (n=9) versus | SD, corn oil (n=9) versus | F344/N, corn oil (n=4) versus |
| **Organ / Morphology** | **F344/N, corn oil or water (n=9)[b]** | **F344/N, corn oil (n=4)[c]** | **F344/N, water (n=5)[d]** |
| Adrenal Medulla, Pheochromocytoma, Benign | < 0.001 *B | 0.051 | 0.002 * |
| Clitoral Gland, Adenoma | < 0.001 *B | < 0.001 *B | 0.066 |
| Clitoral Gland, Carcinoma | 0.003 * | 0.004 * | 0.201 |
| Lung, Alveolar/Bronchiolar, Adenoma | 0.007 * | < 0.001 *B | 0.005 * |
| Mammary Gland, Carcinoma | < 0.001 *B | < 0.001 *B | 0.845 |
| Mammary Gland, Fibroadenoma | < 0.001 *B | < 0.001 *B | 0.079 |
| Multiple Organs, Leukemia, Mononuclear Cell | < 0.001 *B | < 0.001 *B | 0.002 * |
| Multiple Organs, Mesothelioma, Malignant | 0.023 * | 0.033 * | 0.347 |
| Pituitary Gland, Pars Distalis, Adenoma | 0.996 | 0.028 * | < 0.001 *B |
| Pituitary Gland, Pars Intermedia, Adenoma | 0.037 * | 0.093 | 0.938 |
| Skin, Basal Cell, Adenoma | 0.010 * | 0.424 | 0.055 |
| Thyroid Gland, C-Cell Adenoma | < 0.001 *B | < 0.001 *B | 0.728 |
| Thyroid Gland, C-Cell Carcinoma | < 0.001 *B | 0.001 * | 0.924 |
| Thyroid Gland, Follicular Cell Carcinoma | 0.024 * | 0.447 | 0.119 |
| Uterus, Carcinoma | 0.037 * | 0.248 | 0.297 |

[a] An asterisk (*) indicates the p-value is statistically significant at the 0.05 level, in the absence of a correction for multiple testing. An additional superscript B indicates the p-value remains statistically significant after applying a Bonferroni correction for multiple testing, to adjust for 82 tests having been performed.

[b] The first column of p-values corresponds to strain comparisons (Fischer 344/N versus Hsd: Sprague Dawley SD), regardless of vehicle (corn oil or water), among all 18 studies. These p-values were obtained from Table 2.

[c] The second column of p-values corresponds to strain comparisons among the subset of 13 studies having the same vehicle (corn oil).

[d] The third column of p-values corresponds to vehicle comparisons (corn oil versus water) among the nine Fischer 344/N rat studies.

## Table 5

Summary of neoplasms with statistically significant differences in incidence rates between the two strains of control female rats (Fischer 344/N versus Hsd:Sprague Dawley SD) by at least one criterion in our analysis of 18 NTP studies.

| | | | Statistically Significant Strain Difference in Incidence Rates? | | | |
|---|---|---|---|---|---|---|
| | Incidence Rate (%)[a] | | Matched on Vehicle[b] | | Not Matched on Vehicle[b] | |
| Neoplasm | Fischer 344/N | Sprague Dawley | with Bonferroni[c] | without Bonferroni[c] | with Bonferroni[c] | without Bonferroni[c] |
| Clitoral Gland Adenoma | 5.82 | 0.22 | + | + | + | + |
| Mammary Gland Carcinoma | 2.44 | 10.15 | + | + | + | + |
| Mammary Gland Fibroadenoma | 48.44 | 67.44 | + | + | + | + |
| Mononuclear Cell Leukemia | 16.67 | 0.85 | + | + | + | + |
| Thyroid Gland C cell Adenoma | 13.56 | 25.43 | + | + | + | + |
| Lung Alveolar/Bronchiolar Adenoma | 2.00 | 0.42 | + | + | | + |
| Thyroid Gland C cell Carcinoma | 1.33 | 3.85 | | + | + | + |
| Clitoral Gland Carcinoma | 1.34 | 0.00 | | + | | + |
| Malignant Mesothelioma | 0.89 | 0.00 | | + | | + |
| Pituitary Gland Pars Distalis Adenoma | 44.67 | 38.85 | | + | | |
| Adrenal Medulla Benign Pheochromocytoma | 2.89 | 7.23 | | | + | + |
| Pituitary Gland Pars Intermedia Adenoma | 0.00 | 1.06 | | | | + |
| Skin Basal Cell Adenoma | 1.11 | 0.21 | | | | + |
| Thyroid Gland Follicular Cell Carcinoma | 1.11 | 0.21 | | | | + |
| Uterus Carcinoma | 0.22 | 1.27 | | | | + |

[a] Rates obtained from Table 2 and based on nine studies per strain, matched on laboratory (Battelle Columbus), diet (NTP-2000), and route (gavage).

[b] Some analyses further matched on vehicle (corn oil) and some did not (corn oil or water).

[c] The Bonferroni correction adjusts for multiple testing because a separate test was performed for each of the 82 tumor types.