

## Evaluation of Organ Weights for Rodent and Non-Rodent Toxicity Studies: A Review of Regulatory Guidelines and a Survey of Current Practices

BINDHU MICHAEL,<sup>1</sup> BARRY YANO,<sup>2</sup> RANI S. SELLERS,<sup>3</sup> RICK PERRY,<sup>4</sup> DANIEL MORTON,<sup>5</sup> NIGEL ROOME,<sup>6</sup> JULIE K. JOHNSON,<sup>7</sup> AND KEN SCHAFER<sup>8</sup>

<sup>1</sup>*Merck Research Laboratories, West Point, Pennsylvania 19586, USA*

<sup>2</sup>*The Dow Chemical Company, Midland, Michigan 48674, USA*

<sup>3</sup>*Albert Einstein College of Medicine Bronx, New York 10461, USA*

<sup>4</sup>*Wyeth Research, Chazy, New York 12921, USA*

<sup>5</sup>*Pfizer, Inc., Groton, Connecticut 06340, USA*

<sup>6</sup>*Sanofi-Aventis, Porcheville, France*

<sup>7</sup>*Abbott Laboratories, Abbott Park, Illinois 60064, USA*

<sup>8</sup>*Vet Path Services, Inc., Cincinnati, Ohio 45249, USA*

### ABSTRACT

The Society of Toxicologic Pathology convened a working group to evaluate current practices regarding organ weights in toxicology studies. A survey was distributed to pharmaceutical, veterinary, chemical, food/nutritional and consumer product companies in Europe, North America, and Japan. Responses were compiled to identify organs routinely weighed for various study types in rodent and non-rodent species, compare methods of organ weighing, provide perspectives on the value of organ weights and identify the scientist(s) responsible for organ weight data interpretation. Data were evaluated as a whole as well as by industry type and geographic location. Regulatory guidance documents describing organ weighing practices are generally available, however, they differ somewhat dependent on industry type and regulatory agency. While questionnaire respondents unanimously stated that organ weights were a good screening tool to identify treatment-related effects, opinions varied as to which organ weights are most valuable. The liver, kidneys, and testes were commonly weighed and most often considered useful by most respondents. Other organs that break were commonly weighed included brain, adrenal glands, ovaries, thyroid glands, uterus, heart, and spleen. Lungs, lymph nodes, and other sex organs were weighed infrequently in routine studies, but were often weighed in specialized studies such as inhalation, immunotoxicity, and reproduction studies. Organ-to-body weight ratios were commonly calculated and were considered more useful when body weights were affected. Organ to brain weight ratios were calculated by most North American companies, but rarely according to respondents representing veterinary product or European companies. Statistical analyses were generally performed by most respondents. Pathologists performed interpretation of organ weight data for the majority of the industries.

### INTRODUCTION

Organ weight changes have long been accepted as a sensitive indicator of chemically induced changes to organs. In toxicological experiments, comparison of organ weights between treated and untreated groups of animals have conventionally been used to evaluate the toxic effect of the test article (Peters and Boyd, 1966; Pfeiffer, 1968). However, the usefulness of assessing the weights of various individual organs, the manner of presentation of organ weight data, and the value of statistical analyses have also long been topics of discus-

sion. Since organ weight evaluation is an essential part of the toxicologic and risk assessment of drugs, chemicals, biologics, food additives, and medical devices, it is important that the approach to organ weight evaluation and interpretation be done with appropriate scientific rigor and, to the extent possible, consistent with the regulatory guidance.

The Society of Toxicologic Pathology (STP) recognizes certain differences among the regulatory guidelines as well as current practices in organ weights vary widely. Therefore, the STP convened a working group to recommend best practices for the assessment of organ weights in toxicity studies. To aid in making specific recommendations, the working group examined available regulatory guidelines and assessed current work practices regarding the evaluation of organ weights in toxicity studies. This document presents a comprehensive

---

Address correspondence to Sue Pitsch, Society of Toxicologic Pathology, 1821 Michael Faraday Drive, Suite 300, Reston, VA 20190, USA; e-mail: spitsch@toxpath.org

review of regulatory guidelines and the results of a survey conducted to understand the current work practices in the evaluation of organ weights in toxicity studies by pharmaceutical, veterinary, chemical, food/nutritional, and consumer product organizations. The contents of this document represent essentially the views of individual respondents and may not reflect the view of all working group members or the STP.

#### *Review of Regulatory Guidelines on Organ Weight Evaluation in Toxicity Studies*

Various regulatory guidelines differ somewhat in the specific language regarding organ weight recommendations, and a single global standard for organ weights is lacking. The United States Environmental Protection Agency (EPA) guidance, the Redbook from the Center for Food Safety and Applied Nutrition (CFSAN) of the United States Food and Drug Administration (USFDA), Organization for Economic Co-operation and Development (OECD) guidance, Japanese Ministry of Agriculture, Forestry and Fisheries (JMAFF) guidance, European Economic Community (EEC) guidance, Japan's Ministry of Economy, Trade, and Industry (METI), and Japan's Ministry of Health, Labor, and Welfare (MHLW) guidance have listed specific organs to be weighed in a variety of toxicity studies in rodent and non-rodent species (Table 1). The Anticancer Drug Development Guide provides a list of organs generally weighed in toxicity studies for testing oncology products (Roy and Andrews, 2004). In addition, the International Conference on Harmonization (ICH) mentions testicular weights as part of the toxicologic assessment of the male reproductive system in pharmaceutical non-clinical studies, but gives little further guidance for organ weighing. Other guidance documents reviewed did not contain a list of recommended organs to be weighed.

In some cases, the guidance documents specifically mentioned particular organs to be weighed in specialized studies. For example, the European Agency for Evaluation of Medicinal Products (EMA) states that lung should be weighed in inhalation toxicology studies. The EMA and ICH guidances also list specific organs to be weighed for evaluating the immunotoxicity of pharmaceuticals. For immunotoxicity studies, the FDA guidelines for pharmaceutical products only state that "immune system organs" be weighed. The STP has previously made recommendations for weights of lymphoid tissues in a best practices publication (Haley et al., 2005).

#### *Survey Methods*

To document the current practices in the evaluation of organ weights in toxicological studies, a questionnaire was sent to representatives of several pharmaceutical, veterinary, chemical, food/nutritional, and consumer product companies located in Europe, North America, and Japan. This questionnaire solicited responses to determine which organs are weighed in various types of toxicity or safety studies in rodent and non-rodent species, whether paired organs are weighed individually or together, whether organs are weighed before or after fixation, if organ-to-body weight and organ to brain weight ratios are calculated, if weights are statistically analyzed, which personnel interpret organ weight data, and the value of individual organ weights in the assessment of treatment-related effects. The data from completed

questionnaires were tallied, and percentages of responses were calculated for each question according to the respondent's industry type (pharmaceutical, veterinary, chemical, food/nutritional, or consumer product) and geographical location (Europe, Japan, North America, or multinational).

#### RESULTS OF SURVEY

##### *Current Practices in Organ Weight Evaluation in Toxicity Studies*

Approximately 60% of the survey recipients provided responses. The total number and percentage of survey respondents from each industry type were as follows: 31 (91%) pharmaceutical, 3 (9%) veterinary, 4 (12%) chemical, 1 (3%) food/nutritional and 3 (9%) consumer products. Some respondents had product lines in more than one area, and therefore the total percentages were greater than 100%. The total number and percentage of respondents from various geographical locations were as follows: 3 (9%) Europe, 8 (24%) Japan, 10 (29%), North America, and 13 (38%) multinational organizations.

Table 2 summarizes the percentage of survey respondents by type of industry and geographical location for each question in the questionnaire. Because it was anticipated that the respondents would have diverse parameters for their toxicology studies, study types were generally identified generically by study length in the questionnaire.

The use of customized vs. standardized lists for organ weight evaluation varied by industry type and geographical location (Table 2). Customization of the organ weight lists was primarily based on compound activity, study duration, Good Laboratory Practice (GLP) status, and regulatory guidance. Occasionally these lists were based on the species tested or efficacy and previous experience with the test article or its class. All companies, except for 14% of the pharmaceutical companies, weighed paired organs together.

All respondents calculated relative organ weights based on organ-to-body weight ratio (Table 2). Many respondents commented that the organ-to-body weight ratios in combination with absolute organ weight data added value to their interpretation. Most of them cited instances when this ratio was helpful in normalizing the variability due to body weight fluctuations, as in studies where a test article affects the body weight. Some respondents found the organ-to-body weight ratio helpful in normalizing the variability due to nutritional status, and in studies with anti-obesity compounds. To best use the body weights for organ weight evaluation, some respondents stated that it was necessary to take body weight at the time of necropsy and not just a morning weight prior to necropsy.

The percentage of respondents that calculated the organ-to-brain weight ratio varied by industry type and geographical location (Table 2). Many respondents commented that brain weight was considered least prone to body weight changes. Therefore, they found that the organ-to-brain weight ratio was particularly useful in normalizing the variability in organ weight when body weight changes confounded the interpretation of organ weight data. Interestingly, none of the veterinary or European companies calculated organ to brain weight ratios. Some respondents considered organ-to-brain weight ratios unnecessary.

TABLE 1.

	Thyroid with										Seminal Vesicles with				Known target organs		
	Brain	Heart	Liver	Kidneys	Lungs	Spleen	Thymus	Adrenals	Parathyroid	Pituitary	Testes	Epididymides	Prostate	Coagulating glands		Ovaries	Uterus
OPPTS comb repeat dose repro/devel screen, OPPTS and METI 28-day rodent	X	X	X	X	X	X	X	X	X	X	X	X	X				
OPPTS 90-day oral, 90-day dermal, 90-day inhalation	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	
OPPTS 21/28-day dermal	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
OPPTS reproduction & fertility—adults	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
OPPTS reproduction & fertility—weanlings	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	
OPPTS chronic tox, carco, comb chronic tox/carco	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	
OECD comb repeated dose repro/devel screen, 28-day oral rodent	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	
OECD/METI repeated dose 90-day oral tox rodent, OECD 90-day non-rodent	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	
OECD 21/28-day dermal, 90-day dermal, 90-day inhalation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
OECD 28-day or 14-day inhalation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
OECD two-generation reproduction—adults	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	Xf	X	X	Xf
OECD chronic toxicity, combined chronic tox/carcinogenicity	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	
OECD and USEPA OPPTS repro/devel screening	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	
JMAFF 90-day repeated oral tox—rodent	X	X	X	X	X	X	X	Xt	X	X	X	X	X	X	X	X	
JMAFF 90-day repeated oral tox—nonrodent	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
JMAFF 21-day dermal tox, 90-day inhalation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
JMAFF 1-year oral toxicity—rodent	X	X	X	X	X	X	X	Xt	X	X	X	X	X	X	X	X	X
JMAFF 1-year oral toxicity—nonrodent	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
JMAFF 1-year repeated dose oral tox/carco comb	X	X	X	X	X	X	X	Xt	X	X	X	X	X	X	X	X	X
JMAFF reproductive—adults	X	X	X	X	X	X	X	Xt	X	X	X	X	X	X	X	X	X
JMAFF reproductive—offspring	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	X
EEC subchronic oral tox—rodents and non-rodents	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	X
EEC subchronic dermal tox—rodent	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
EEC subchronic inhalation tox—rodent	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
EEC chronic tox, combined chronic tox/carco in rodents	X	X	X	X	X	X	X	Xnr	X	X	X	X	X	X	X	X	X
CFSAN Redbook short term, subchronic, carcinogenicity, 1 year tox	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CFSAN Redbook reproductive—adult F0, F1 parents	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CFSAN Redbook reproductive—weanlings F1 and F2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
MHLW Japan guidelines (absolute & rel. to BW)	Xa	Xa	Xa	Xa	Xo	Xa	Xo	Xto	Xa	Xa	Xa	Xa	Xso	Xa	Xo	Xo	Xo
METI Japan guidelines: chronic toxicity	X	X	X	X	X	X	X	X	X	X	X	X	Xs	X	X	X	X
Anticancer drug development guide	X	X	X	X	X	X	X	Xt	X	X	X	X	Xs	X	X	X	X
ICH S8/EMEA immunotox guidance, STP immunotox position paper	X	X	X	X	X	X	X	Xt	X	X	X	X	Xs	X	X	X	X

ICH immunotoxicity guidelines include lymph nodes as optional while EMEA guidance include draining and distant lymph node. In MHLW guidelines, Xa indicate "always" and Xo indicates "often." Xf denotes including their fluids. Cervix and vagina are also included in anticancer drug development guide. Cervix is also included in EEC teratogenicity guidelines for rodents and non-rodents. OECD prenatal developmental toxicity guidelines indicate weighing gravid uteri including cervix. Xt denotes only thyroid is weighed. Xnr means thyroid is weighed with parathyroid only in non-rodents. SO denotes seminal vesicle is often weighed. S denotes only seminal vesicle is weighed. No organs were listed for weights in the following guidelines: USEPA OPPTS prenatal development, OECD teratogenicity, OECD carcinogenicity, JMAFF teratogenicity, EEC carcinogenicity, EEC one-generation reproduction, and EEC two-generation reproduction.

TABLE 2.

	Pharma	Veterinary	Chemical	Food/ Nutritional	Consumer Products	Europe	Japan	North America	Multinational
Total respondents (%)	91	9	12	3	9	9	24	29	38
Standard list of organs weighed	83	33	50	100	100	33	75	90	92
Customized list of organs weighed	33	67	75	100	33	67	25	20	50
Customize based on:									
Study length	33	33	25	0	0	33	38	20	25
GLP status	17	0	0	0	0	67	13	0	17
Compound activity	70	33	75	100	67	67	63	50	67
Regulatory Guidelines	17	33	50	0	0	67	25	10	17
Species	3	0	0	0	0	0	0	0	0
Efficacy	3	0	0	0	0	0	0	0	0
History of compound	3	0	0	0	0	0	0	0	8
Paired organs weighed together <sup>a</sup>	86	100	100	100	100	100	100	100	100
Paired organs weighed individually	14	0	0	0	0	0	75	0	0
Organs are weighed after fixation	50	67	75	100	67	100	63	67	23
Organs are not weighed after fixation	50	33	25	0	33	0	38	33	77
Calculate organ weight to body weight	100	100	100	100	100	100	100	100	100
Calculate organ weight to brain weight	43	0	25	100	33	0	38	56	46
Do not calculate organ weight to brain weight	57	100	75	0	67	100	63	44	54
Statistical analysis for routine toxicity studies?									
Yes	90	67	100	100	67	100	86	56	92
No	10	33	0	0	33	0	14	33	0
N/A	0	0	0	0	0	0	0	11	8
Statistical analysis for carcinogenicity studies?									
Yes	29	67	100	100	33	33	67	38	15
No	46	33	0	0	33	33	33	50	62
N/A	18	0	0	0	33	33	0	13	23
Comparison to concurrent controls	74	33	50	0	33	67	50	56	100
Comparison to concurrent and historic controls	26	67	50	100	67	33	50	44	0
Pathologist interprets the organ weight data	74	67	75	100	100	0	63	100	82
Pathologist do not interpret the organ weight data	19	33	25	0	0	100	38	0	0

<sup>a</sup>Paired organs were weighed together for preliminary tests and individually for main tests by 33% chemical industry and 25% Japanese respondents. Kidneys were weighed individually while gonads were weighed together by 8% multinational respondents.

Statistical analyses of organ weight data were done in routine toxicity studies by all the chemical, food/nutritional, and European industry respondents as well as the majority of the other respondents. Some respondents commented that a pattern of relatively similar changes in absolute organ weights, organ-to-body weight ratios, and organ to brain weight ratios were needed to attribute the organ weight changes to a test article effect. They believed that statistical analyses did not always enhance the understanding of these effects and could be misleading.

The percentage of respondents who weigh organs in carcinogenicity studies and/or perform statistical analyses of organ weight data in carcinogenicity studies varied among the industry types and geographical locations (Table 3). All of the chemical and food/nutritional companies weighed organs in carcinogenicity studies and subjected them to statistical analysis. Reasons stated by the respondents for not performing statistical analyses of organ weights in carcinogenicity studies included the increased variability in organ weights in long-term studies, and the presence of tumors that may confound results. Further reasons cited included the lack of any additional insight into test article effects, especially if sufficient number of animals were used in shorter-term studies.

Comparison of organ weights from test article treated groups to concurrent controls or both concurrent and historical controls varied among the industry types and geographical locations (Table 2). Many respondents routinely compared the organ weights from test article treated groups to concurrent controls and only utilized the historical controls in special situations, such as when the findings were equivocal, the concurrent control data were inadequate, or

when aberrant concurrent control values skewed the mean. They found historical controls useful to explain outliers and inconsequential or irrelevant statistical flags and in situations where the concurrent controls had limited variability or were outside of the historical range. Some respondents commented that the historical control data were particularly useful in non-rodent studies.

The pathologist interpreted the organ weight data in all of the food/nutritional, consumer product and North America based companies and the majority of the other companies except for European companies, where none responded that the pathologist interpreted organ weights. Additionally, 6% of pharmaceutical and 18% of multinational industries indicated that there was site-to-site variation in whether a pathologist interpreted the organ weight data. Further, 6% of pharmaceutical and 25% of Japanese companies specified that the study director interpreted the organ weight data in their toxicity studies.

The specific organs weighed by the respondents in each type of industry and geographical location in both rodent and non-rodent toxicity studies are summarized in Table 3. Table 3 is not all-inclusive of the organs weighed as some of the respondents occasionally weighed additional organs in certain types of studies. For example, tumors were weighed in rodent carcinogenicity studies by 1 of the Japanese respondents (17%); cecal weights were evaluated by 1 respondent from the chemical industry (25%) and 1 from Japanese industry (17%); gall bladder weights were measured in large animal GLP subchronic toxicity studies by 33% of European respondents (1 respondent), 4% of pharmaceutical respondents (1 respondent), 33% veterinary





TABLE 4.—Organ weights contributing most value to studies.

Ranking	Pharmaceutical	Veterinary	Chemical	Food/Nutritional/Consumer
1	Liver	Liver	Liver	Liver
2	Testes	Kidneys	Kidneys	Kidneys
3	Kidneys	Testes	Testes	Testes
4	Adrenal	Adrenal	Spleen	Epididymides
5	Thymus	Thyroid/Thyroid parathyroid	Thyroid/Thyroid parathyroid	Thymus
6	Brain	Brain	Heart	Thyroid/thyroid parathyroid
7	Spleen	Heart	Thymus	Spleen
8	Ovary	Spleen	Brain	Lungs
9	Heart	Lung	Epididymides	Adrenal
10	Prostate	Ovaries	Lung	
11	Thyroid or Thyroid/ Parathyroid		Ovaries	
12	Lung			
13	Uterus			
14	Pituitary			
15	Epididymides			
16	Lymph nodes			

respondents (1 respondent), and 25% of chemical industry respondents (1 respondent). In addition, pancreas was occasionally weighed: among the pharmaceutical industry as a whole, the pancreas was weighed by 4%, 3%, and 5% of respondents in large animal GLP subchronic toxicity studies, rodent GLP subchronic toxicity studies, and large animal non-GLP repeat dose toxicity studies, respectively (1 respondent in each category); among Japanese respondents, 13% weighed pancreas in large animal non-GLP repeat dose toxicity studies and large animal GLP subchronic toxicity studies (1 respondent); among multinational respondents, 1 weighed pancreas in rodent GLP subchronic toxicity studies.

#### *Additional Organs Weighed in Special Studies*

Several respondents weighed additional organs in specific types of studies: draining lymph nodes were weighed in vaccine studies, while lungs, nasal turbinates, draining lymph nodes, and bronchial lymph nodes were weighed in inhalation studies. Spleen, lymph nodes (such as popliteal, mesenteric, mandibular or unspecified peripheral nodes), and thymus were weighed in immunotoxicity studies. In developmental toxicity studies, many respondents weighed additional tissues depending on the anticipated toxicity. These included testes, prostate, epididymides, seminal vesicles, ovaries, uterus (alone or with cervix), placenta, and fetuses.

Some respondents specified that testes, epididymides, prostate, and seminal vesicles were weighed in male rat fertility studies, while brain, testes, epididymides, prostate, seminal vesicles, ovaries, and uterus were weighed in pre- and postnatal studies. In juvenile toxicity studies, some respondents only weighed testes, epididymides, prostate, ovaries, and uterus, while others weighed adrenal glands, brain, heart, kidneys, liver, ovaries, pituitary, prostate, spleen, testes, thyroid/parathyroid, and thymus. Some respondents weighed the cecum in antibiotic studies, rodent GLP subchronic studies, or rodent range finding studies to support carcinogenicity studies.

#### *The Value of Organ Weight Data in Toxicity Studies*

Respondents were asked to comment on the organ weights that contributed most and least to data interpretation. Organ weights that were considered most valuable are summarized by industry in Table 4.

Liver weight was considered useful by 81% of the pharmaceutical industry respondents and 100% of the respondents in other industries. Factors cited for its usefulness included: its sensitivity to predict toxicity; it is a frequent target in toxicity studies; it is useful to evaluate/support diagnosis of hepatocellular hypertrophy from hepatic enzyme induction, peroxisome proliferation or lipidosis; it is often reflective of physiologic perturbations and metabolism; it correlates well with histopathological changes; there is little animal-to-animal variability; there is available historical control range data; and it is important as the primary detoxification organ.

Kidney was considered useful by 58% of pharmaceutical, 67% of consumer product industry respondents and 100% of the respondents in other industries. Reasons cited for the usefulness of weighing kidneys in toxicity studies included: its sensitivity to predict toxicity, enzyme induction, physiologic perturbations and acute injury; it is frequently a target organ of toxicity; it correlates well with histopathological changes; there is little interanimal variability; and historical control range data is available. Some respondents believed that weight changes were more sensitive than histopathologic changes in the kidney. Although 3% of pharmaceutical and 33% of veterinary and consumer product industry respondents commented that kidney weights were not particularly useful, they did not cite any explanations.

Testicular weights were considered valuable by 74% percent of pharmaceutical, 67% of veterinary and consumer product, 75% of chemical, and 100% of food/nutritional product industry respondents in sexually mature animals. The main reasons cited for the usefulness of testicular weights in toxicity studies were: their sensitivity to toxicity due to perturbations in rapidly dividing cells, perturbations in physiology, or perturbations in hormones, and weights are useful to determine maturity. Other reasons cited included aiding in the identification of enzyme induction; they correlate well with histopathologic changes; and testicular weights are useful in establishing a No-Observed-Effect-Level (NOEL), even in the absence of a morphologic correlate. One pharmaceutical industry respondent stated that testicular weights were unreliable in immature non-rodents, particularly non-human primates.

Adrenal gland weights were considered to be a sensitive predictor of toxicity and stress by 45% of pharmaceutical, 67% of veterinary, and 33% of consumer product industry

respondents and weights correlated well with histopathological hypertrophy/hyperplasia. Other reasons cited for the usefulness of adrenal gland weights include low interanimal variability and the availability of historical control range data. In addition, some remarked that, because of stress induced hypertrophy, adrenal glands were useful in the interpretation of whether lymphoid changes were stress-related or due to immunotoxicity. However, 6% of pharmaceutical respondents, 50% of chemical industry and 67% of consumer product respondents found adrenal gland weights to be poor predictors of toxicity because they are often affected by stress and may not correlate with other primary treatment-related findings. Additionally, due to the presence of extraneous tissue and difficulty in trimming, 33% of veterinary and 25% of chemical industry respondents found adrenal glands weights to be unreliable in mice.

Although some respondents commented that it was uncommon to see a change in brain weight due to toxicity, 39% of pharmaceutical, 33% of veterinary, and 50% of chemical industry respondents considered brain weight useful to calculate organ to brain weight ratios for normalizing any variability between groups, particularly when body weight changes were present.

Thymic weights were considered useful by 45% of pharmaceutical, 50% of chemical, 100% food/nutritional, and 67% of consumer product industry respondents and spleen weights were considered valuable in 36% of pharmaceutical, 33% of veterinary, 75% of chemical, 100% of food/nutritional, and 67% of consumer product industry respondents. Thymus and spleen weights were considered to be sensitive indicators of immunotoxicities (immune stimulation or depletion), stress, and physiologic perturbations. Furthermore, histopathologic changes in the spleen and thymus were thought to correlate to organ weight changes. Conversely, the thymus was believed to have little value in 26% pharmaceutical, 33% of veterinary, and 25% of chemical industry respondents and spleen weights were believed to be of little value by 39% of pharmaceutical, 33% of veterinary and 25% of chemical industry respondents.

Explanations for why thymus weights were considered to have limited value included variability from factors such as dissection technique and age-related involution (especially in non-rodents), and stress-related effects. Spleen weights were deemed of limited value due to interanimal variability; stress-related effects; and physiologic factors unrelated to treatment, such as euthanasia-associated splenic congestion in non-rodents (dogs in particular). In addition, these respondents thought that spleen and thymus weight changes did not often correlate with histopathologic findings, and histopathology was considered a more sensitive indicator of test article effects.

Heart weight was considered of value by 32% of pharmaceutical, 33% of veterinary, and 50% of chemical industry respondents because of its correlative nature with hypertrophy, its sensitivity to predict toxicity, its limited interanimal variability, and the availability of historical control range data. However, 10% of pharmaceutical, 50% of chemical industry, 100% of food/nutritional, and 67% of consumer product industry respondents considered heart weights of limited value because of variability due to small size or physiologic factors unrelated to treatment and the lack of cor-

relation with other findings. It was commented that decreased heart weight had no microscopic correlate and were often associated with body weight decreases secondary to treatment.

Ovarian weights were considered valuable by 35% of pharmaceutical, 33% of veterinary, and 25% of chemical industry respondents, particularly in sexually mature animals because they could indicate ovarian dysfunction and help in establishing a NOEL even in the absence of a morphologic correlate. However, ovarian weights were considered of little value by 19% of pharmaceutical, 50% of chemical, 100% of food/nutritional, 67% of consumer product industry respondents. Uterine weights were considered valuable by 19% of pharmaceutical respondents, but were considered of little value by 29% of pharmaceutical, 33% of veterinary, 75% of chemical, 100% of food/nutritional, 67% of consumer product industry respondents. Factors that diminished the usefulness of uterine and ovarian weights were stated to be variability due to its small size and technical issues with consistent dissection; physiologic factors unrelated to treatment (such as estrus cycle); and the relative infrequency of these organs as target tissues. In addition, respondents stated that ovarian and uterine weights were often decreased with a decrease in body weight due to systemic (nonspecific) toxicity. Some respondents stated that reproductive tract histopathology was a more sensitive indicator of toxicity.

Thyroid or combined thyroid/parathyroid gland weights were considered of value by 29% of pharmaceutical, 67% of veterinary, 75% of chemical, 100% of food/nutritional, and 67% of consumer product industry respondents. Reasons cited for their usefulness were low inter-animal variability; availability of historical control data; good correlation to histopathologic findings; and sensitivity as a predictor of toxicity. Many respondents stated that thyroid gland weight was sensitive to decreased thyroxin levels and thus could be an indicator of hepatic enzyme induction in conjunction with liver weight and morphology. Thyroid gland weight was stated to have limited value by 13% of pharmaceutical, 33% of veterinary, and 25% of chemical industry respondents due to variability due to its small size and technical issues with consistent dissection (especially in rats and mice); physiologic factors unrelated to treatment; and because thyroid weight increases often reflected hepatic enzyme induction. Some of these respondents thought that histopathology was more sensitive for identification of toxic effects than weight changes.

Lung weight was often considered useful only in inhalation studies. Nineteen percent of pharmaceutical, 33% of veterinary, 50% of chemical, 100% of food/nutritional, and 67% of consumer product industry respondents commented that lung weights were useful. They commented that lung weight changes correlated with histopathologic changes and were often related to toxicity or enzyme induction. Nineteen percent of pharmaceutical industry respondents considered lung weights less useful due to variability from poor or indistinct demarcations for trimming off the airways; a lower frequency of finding weight changes that correlated with toxicity; and less sensitivity to predict toxicity compared to histopathology.

Epididymal weight was considered useful by 13% of pharmaceutical, 50% of chemical, 100% of food/nutritional, and 67% of consumer product industry respondents because it

was considered to be a common target organ and weight changes often correlated well with histopathologic findings. However, 3% of pharmaceutical industry respondents commented that it was unreliable in immature non-rodents, particularly non-human primates. Prostate gland weights were identified as a sensitive indicator of hormone perturbations in sexually mature animals by 29% of pharmaceutical industry respondents. Nineteen percent of pharmaceutical industry respondents considered prostate weights less useful due to its small size and technical issues with consistent dissection (poor or indistinct demarcation); physiologic factors unrelated to treatment; poor correlation of weight changes with histopathologic findings; or non-specific reductions in weight due to decreased body weight.

Pituitary weights were considered valuable as a sensitive indicator of toxicity by 13% of pharmaceutical industry respondents. The value as an indicator of toxicity was facilitated by the availability of historical control range data and low interanimal variability. Thirty-nine percent of pharmaceutical, 67% of veterinary, 75% of chemical, 100% of food/nutritional, and 100% of consumer product industry respondents deemed pituitary weights less useful because of variability due to its small size and technical issues with inconsistent dissection (esp. in rats and mice); poor correlation with histopathological findings; and physiologic factors unrelated to treatment. These respondents thought that histopathology was more sensitive for identifying treatment-related changes than were organ weights.

Lymph node weights were identified as helpful in identifying immunomodulatory effects by 6% of pharmaceutical industry respondents. However, 10% of pharmaceutical industry respondents considered lymph node weights of little value due to the variability between animals and uncertainty about the number of mesenteric lymph nodes to weigh.

Respondents made the following suggestions to minimize the variability in the organ weights and thus enhance the usefulness of organ weight evaluation in toxicity studies: randomization of the necropsy order; randomization of prosector during necropsy to avoid bias; and sufficient sample size (e.g., 10 per sex per group in rats).

#### CONCLUSIONS

The current regulatory guidance available to different industries conducting toxicology studies do vary with regard to specific language about the collection of organ weight data. In general, the respondents to the STP survey considered organ weights to be a useful screening tool to characterize test article-related effects in general toxicity studies. However, the opinions varied widely as to which organ weights are most useful. In addition, some of the respondents remarked that organ weight changes in and of themselves were not necessarily toxic effects. Some emphasized

that the organ weight data should be assessed in the context of the entire study. The context include consideration of body weight changes, pharmacologic action of test article, clinical pathology data, knowledge of the animal's fasted state or if exsanguinated, as well as macroscopic and microscopic findings.

#### REFERENCES

- Peters, J. M., and Boyd, E. M. (1966). Organ weights and water levels of the rat following reduced food intake. *J Nutr* **90**(4), 354–60.
- Pfeiffer, C. J. (1968). A mathematical evaluation of the thymic weight parameter. *Toxicol Appl Pharmacol* **13**(2), 220–7.
- Roy, D., and Andrews, P. A. Nonclinical testing: from theory to practice. (2004). In: *Anticancer drug development guide* (B. A. Teicher and P. A. Andrews, eds.), pp. 306. Humana Press, Totowa, New Jersey.
- United States Food and Drug Administration (USFDA) Center for Food Safety and Applied Nutrition (CFSAN) Office of Food Additive Safety Redbook Toxicological Principles for the Safety Assessment of Food Ingredients, (2000). <http://www.cfsan.fda.gov/~redbook/red-toca.html>
- European Economic Community (EEC), European Chemicals Bureau, Toxicology and Chemical Substances Guidelines. <http://ecb.jrc.it/testing-methods/>
- International Conference of Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) Topic S5, Detection of Toxicity to Reproduction for Medicinal Products and Toxicity to Male Fertility, 2005. <http://www.ich.org/cache/compo/276-254-1.html>
- International Conference of Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) Topic S8 Immunotoxicity Studies for Human Pharmaceuticals, Note for Guidance on Immunotoxicity Studies for Human Pharmaceuticals, 2006. <http://www.ich.org/cache/compo/276-254-1.html>
- United States Environmental Protection Agency (EPA), Office of Prevention, Pesticides and Toxic Substances (OPPTS), Health Effects Test Guidelines. <http://www.epa.gov/opptsfrs/home/guidelin.htm>
- Haley, P., Perry, R., Ennulat, D., Frame, S., Johnson, C., Lapointe, J.-M., Nyska, A., Snyder, P. W., Walker, D., and Walter, G. (2005). *Toxicol Pathol* **33**, 404–7.
- Organization for Economic Cooperation and Development (OECD), Environment Directorate, Guideline for the Testing of Chemicals. [http://www.oecd.org/department/0,2688,en\\_2649\\_34377\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/department/0,2688,en_2649_34377_1_1_1_1_1,00.html)
- The European Agency for the Evaluation of Medicinal Products Committee for Proprietary Medicinal Products, Note for Guidance on Repeated Dose Toxicity, 2000. <http://www.emea.eu.int/pdfs/human/swp/104299en.pdf>
- Japanese Ministry of Agriculture, Forestry and Fisheries, Agricultural Chemical Inspection Station, Data Requirements for Supporting Registration of Pesticides Annex (Guidelines for Preparation of Study Results Submitted When Applying for Registration of Agricultural Chemicals). <http://www.acis.go.jp/eng/shinsei/annex.htm>
- Japanese Ministry of Economy, Trade, and Industry, Policy Information for Pre-manufacturing Evaluation System for New Chemical Substances [http://www.meti.go.jp/english/information/data/chemical\\_substances03\\_1.html](http://www.meti.go.jp/english/information/data/chemical_substances03_1.html)
- Guidelines for Toxicity Studies of Drugs Manual, Ministry of Health, Labour and Welfare, Pharmaceutical Affairs Bureau, New Drugs Division, 1990.