Exponent®

TECHNICAL MEMORANDUM

TO:	Dr. Christian Strupp, ADAMA Deutschland GmbH
	Dr. Loïc Bretesché, ADAMA France SAS
FROM:	John M. DeSesso, Ph.D., Fellow ATS, Exponent, Inc.
	Amy L. Williams, Ph.D., DABT, Exponent, Inc.
DATE:	13 November 2015
PROJECT:	1500532.000, Task B0T0
SPONSOR REFERENCE:	90019520
SUBJECT:	Expert Opinion Regarding Male Reproductive Toxicity Study of Sarkar et al., 1997

Isoproturon is under consideration for classification regarding reproductive toxicity in the European Union (EU) based in part on findings reported in a study published by Sarkar et al. (1997). The purpose of this technical memorandum is to provide an expert opinion regarding the relevance of findings of male reproductive toxicity as reported in this study. Results from a related published study (Sarkar et al., 1995) on the subacute toxicity of isoproturon in male rats are also considered in this analysis. For the purposes of this assessment, the two studies by Sarkar et al. were reviewed. Additional data available in the published literature and in the draft renewal toxicological assessment report for isoproturon (2015; provided to us by the client) were also considered.

In the study by Sarkar et al. (1997), adult male albino rats are treated with 0, 200, 400 or 800 mg/kg/day of isoproturon by oral gavage for 6 days per week for a total of 10 weeks. The vehicle was reported to be "refined ground-nut oil." Various findings of male reproductive toxicity were reported at the high dose of 800 mg/kg/day, including reduced epididymal sperm counts, reduced % motile sperm, increased % abnormal sperm, reduced testicular seminiferous tubule mean diameter; and reduced % tubules with evidence of spermatogenesis. An increase in the % of damaged tubules is also reported at both 400 and 800 mg/kg/day. Our analysis finds that there are significant methodological flaws with this study that negate one's ability to draw any conclusions from the data. These issues include the use of apparently malnourished animals in the study; incorrect methods for the preparation and fixation of testicular tissues; and the failure to use appropriate stereological methods. We further note a lack of sufficient

methodological detail on how the study was conducted and the use of too few animals or samples from which to draw any firm conclusions. Each of these elements is discussed below in detail.

The use of malnourished animals

The specific stain of rats used in this study and their age are not reported. Rather, the study reports that "adult male albino rats weighing between 95-110 g body weight were used" (Sarkar et al., 1997). Other details regarding the animals and experimental design are provided in an earlier study (Sarkar et al., 1995), which reports that the animals were bred on site at the Indian Veterinary Research Institute (IVRI). Specific information regarding the growth and development of the rats used in this study or at the IVRI are not readily available. However, Sprague Dawley and Wistar are both strains of albino rats about which substantial information is available regarding their growth and development. A comparison of data available for these two strains of rats with data reported for the control rats in the two Sarkar studies shows that the rats used in the study of concern were significantly underweight.

Figures 1 and 2 below show the growth curves for Sprague Dawley and Wistar rats, respectively. These data are from Charles River Laboratories in the United States, a supplier of rats and other animals for research studies. These data show that male Sprague Dawley rats are typically around 95-110 grams at approximately 5 weeks of age; male Wistar rats are typically in this weight range around 4 weeks of age. These two strains of rats typically reach puberty around 6 weeks of age (based on the time of preputial separation) and usually weigh around 200 grams at puberty (EPA, 2009). It is sometime shortly after the animals have reached puberty that they would be considered adults. Although Sarkar et al. (1997) describe the rats used in their studies as being adult, they weighed considerably less at study initiation (~100 grams less) than one would expect for an adult male rat.

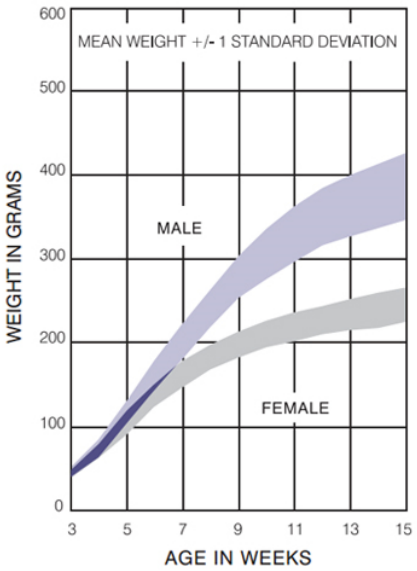

Figure 1. Growth Curve for Sprague Dawley Rats (from http://www.criver.com/products-services/basic-research/find-a-model/sprague-dawley-rat).

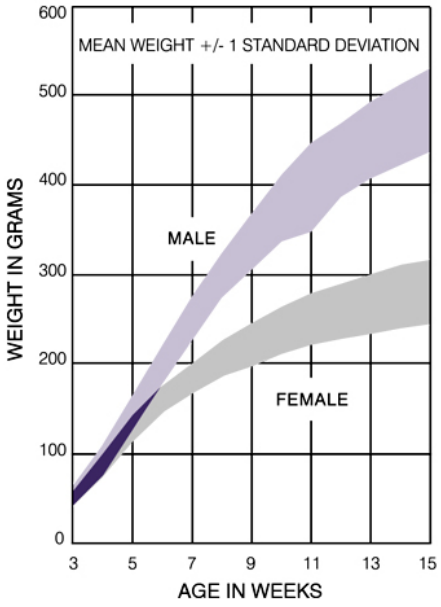

Figure 2. Growth Curve for Wistar Rats (from http://www.criver.com/products-services/basic-research/find-a-model/wistar-rat).

The body weights of the animals at the end of the 10-week treatment period are reported in the earlier study (Sarkar et al., 1995). The controls only gained ~79 grams in this time frame. This is significantly less than one would expect based on the data provided by Charles River for Sprague Dawley and Wistar rats. Based on the growth curves provided above, between 5 and 15 weeks of age, these rats typically gain ~280 grams and ~330 grams, respectively. It should be further noted that the rats in the mid-dose group (400 mg/kg/day) gained slightly less than controls (~69 grams) and rats in the high dose group (800 mg/kg/day) only gained 25 grams over the 10-week treatment period. The minimal weight gain in the high dose group indicates substantial systemic toxicity.

Although information regarding food consumption is not available from the Sarkar studies, the minimal weight gain of all of the rats (including controls) suggests that the animals may suffer from significant nutritional deficits compared to rats typically used in research studies. This conclusion is further supported by description of the diet given in Sarkar et al. (1995):

> *Rat diet was diet composed of crushed wheat 60%, crushed maize 30%, wheat bran 7.5%, mineral mixture 1.25% and common salt 1.25%. In addition, milk powder (Anik-spray) was added to the above mixture @ 2 g/rat.*

This diet is likely inadequate by today's standards. The nutritional deficits associated with the diet provided in this study further compound the systemic toxicity observed at the highest dose (800 mg/kg/day).

Finally, it must be emphasized that the minimal weight gain of the rats in this study is likely to have affected many, if not all, of the parameters assessed. Rehm et al. (2008) previously showed that food restriction affected the absolute weight of the epididymis and was associated with

spermatocyte degeneration in male Sprague Dawley rats; food restriction also caused reductions in serum testosterone levels. Thus, it is possible that the reductions in the enzymes involved in testosterone production observed in Sarkar et al. (1997) may be due to the nutritional deficits and minimal weight gains observed in the study. Food restriction has also been shown to affect sperm numbers (Brinkworth et al., 1992). In summary, the nutritional deficiencies and lack of adequate weight gains observed in all treatment groups (including control) in the Sarkar studies negates one's ability to draw conclusions regarding the effects of isoproturon on male reproductive endpoints.

Incorrect methods for the preparation and fixation of testicular tissues

Excellent fixation of the testes is imperative when assessing potential microscopic damage to the seminiferous tubules or interstitial areas. There are two methods which should be employed for appropriate fixation of the testis. The first method involves preparation of the testicular tissue prior to fixation. Because the testes are encased in a thick capsule (tunica albuginea), fixation by immersion (placing the entire testis in a volume of fixative) is not recommended unless the testis is sectioned or diced prior to entering the fixative. A less effective method, but one which is still superior to immersion of the full testis, involves piercing the testis multiple times with a scalpel blade to allow access of the fixative to the labile tissues within the tunica albuginea. The second method involves selection of an appropriate, fast penetrating fixative. Rapid penetration of fixative is enabled by inclusion of acetic acid in a fixative such as Bouin's solution and Davidson's solution. One fixative that regulatory agencies specifically mention should be avoided is formalin (see discussion in Latendresse et al., 2002).

The Sarkar et al. (1997) study failed to utilize either of these methods. The paper provides no information concerning sectioning, dicing or incising of the testes prior to immersion in fixative. This kind of detail would not be omitted had it been performed. The fixative used was 10% formalin, the one fixative that regulatory agencies recommend avoiding due to poor fixation.

The impact of poor penetration into the testis is that tissue in the center of a testis may not be preserved before autolysis begins. If the tissues used for examination in the treated testes came from the center of the testis while control testes were examined near the periphery, there may be confounding due to the rate of penetration of the fixative. As discussed further below, the authors provide no records concerning the location within the testis for each of the fields that were evaluated. A common (subconscious) bias for microscopists is to select control tissue that has good fixation and staining properties to photograph, whereas the selection is not so strong when evaluating tissues from treated animals. This may have occurred in the Sarkar study.

The failure to use appropriate stereological methods

Microscopy is a powerful tool for studying tissue organization and the structure of cells and for observing changes in the size or distribution of tissue layers and cells that may be associated with alterations in function. Nevertheless, without following a careful design, it is difficult to

impartially quantify these changes. There are several reasons for this. Importantly, morphological alterations occur within a normal physiological range; thus, excursions beyond the normal range – as may occur under adverse conditions – are often difficult to recognize unless they are extreme. Additionally, photomicrographs of histological sections are two-dimensional representations of tissues that have complex (and often asymmetric), three-dimensional anatomy. When comparing experimental versus control tissues, the situation is further confounded by the small amounts of tissue that are actually fixed, blocked, and sectioned relative to the total amount of tissue that is present in the organ under study. It must also be recalled that the fields of view evaluated have been enlarged 400 or more times. This results in a condition wherein the actual amount of tissue viewed is small compared to the entire organ, but the observer has a false sense of having examined a large amount of tissue. A method to ensure that the observations reported are unbiased and capable of providing information about the entire organ under study is stereologic morphometry.

Ewald Weibel, a leader in this field, wrote several review papers and book chapters in the late 1960's and 1970's that describe the principles of stereologic morphometry and how to apply these to electron microscopy (*e.g.*, Weibel et al, 1966; Weibel, 1969; Weibel and Bolender, 1973), which ensure that the data collected are representative of the entire organ in a given animal and have been selected in a manner that is objective and unbiased, and thus allow for quantification. The breadth of the principles discussed in these papers is extensive, but the most important of these with regard to Sarkar et al. (1997) are summarized briefly below. It is clear that, although they claimed to have performed stereologic analyses, Sarkar et al. failed to do so.

- **Tissue sampling**
    - Appropriate number of animals per group
    - Appropriate number of blocks to be prepared per tissue sample
    - Methods to ensure adequate diversity in locations viewed by microscopy for each sample examined from a given organ (using only areas with excellent or good fixation/staining biases the results!)
    - Methods to ensure comparability between experimental and control tissue samples
- **Methods for analysis of micrographs**
    - Analysts should be blinded to treatment and (if possible) to hypothesis
    - Numbers of sections per block usually should be no more than one
    - Magnification and location of recorded fields should be equivalent across all treatment and control groups
    - Stereologic methods for data acquisition
        - Appropriate distribution of blocks to be cut
        - Appropriate methods for random assessment of fields to be examined
        - Appropriate methods to address questions related to changes cell appearance, volume, or number

- Statistical analyses
  - Should be used *a priori* to determine the numbers of animals, tissues, blocks and sections to be analyzed as well as the design of the grid to be used for acquiring data
  - Methods for data analysis should be selected based upon hypothesis to be tested and magnitude of difference to be considered significant

Stereologic methods are exacting and labor intensive, requiring copious samples from many animals. Relatively few sections are actually analyzed per animal, however, and often only a single section per tissue block is used for analysis. As arduous as the methods may seem, these procedures ultimately provide the researcher with an unbiased data set that can then be subjected to rigorous statistical analysis to support quantifiable results for observations that are modulations of normal variation in cellular physiology.

Although Sarkar et al. (1997) attempts to address changes in cellular structure and numbers by what they call 'stereologic analysis', the study does not describe methods that adhere to the principles for stereologic morphometry, as outlined above. To illustrate, the methodological issues regarding the micrscopy described for Sarkar et al. (1997) will be discussed in some detail below.

The Sarkar et al. (1997) study does not describe how the microscopic fields of the seminiferous tubules were selected for analysis. In qualitative morphological studies, microscopists look for sections that are well-fixed and well-stained. These types of micrographs are good for publication purposes, but they are not useful in morphometric analyses. Poorly fixed regions must also be examined as these may have been adversely affected by treatment, and could provide valuable information. A plan for selecting the location of micrographic fields to be photographed and a design for acquiring data from the micrographs should be decided upon *a priori* and should be included in the study's methods description. Such procedures, however, are not described in Sarkar et al. (1997).

Study of sections of the testes is complicated by the stages of the seminiferous tubules. Rats have 14 (I - XIV) stages based on characteristics of the spermatogenic cycle (e.g., Creasy, 1997). These stages repeat themselves as one moves along the length of the seminiferous tubule. The number and maturity of spermatids change for each stage. It is inappropriate to compare spermatid numbers within seminiferous tubules between treated and control groups unless the tubules are at the same stage. There is no indication that the Sarkar group staged the tubules within which they recorded the % of tubules with spermatogenesis. In addition, the nutritional status of the animal affects the quality of sperm production (Rehm et al., 2008), which further complicates the interpretation of the results of the Sarkar study.

The methods reported for quantitation are sparse. Although not given in the methods description, from various footnotes in the tables, it appears that 3-5 testes per dose group were examined histologically. For each testis, 3-5 fields were examined. It is not clear in which region of the testis the histological sections were taken (or if the sections for each testis were taken from the same location across animals and groups). Sections of the testis are large compared to a microscopic field of view. It is not explained if the fields from various groups were taken at the same location (recall that tissue penetration is slow for formalin-based fixatives, so if the locations for treated groups were deeper than those for controls, there would be a possibility of autolysis in the treated tissues that could be interpreted as being compound-related). Were the 3-5 field examined for each testis from the same histologic section? If not, how many sections intervened between the sections selected for examination? Most importantly, regardless of the previous answers, were the methods for acquiring the fields to be examined the same across all groups? If not, the data acquired are not sufficiently similar to allow unbiased analysis.

These are serious shortcomings in methodology, especially when the findings derived from the observations made on a small number of histological sections are meant to determine a potential change in cellular characteristics or numbers that regularly fluctuate temporally within a normal range within healthy seminiferous tubules. Value judgments regarding "damaged seminiferous tubules" may have been biased by the preceding shortcomings. In addition, the authors reported the findings as percentages and the denominators differed between dose groups making this even more unreliable as an unbiased value.

Insufficient methodological detail and the use of too few animals or samples in sperm motility and enzyme activity studies

The numerous methodological deficiencies in the Sarkar et al. (1997) study with regard to the animal husbandry, testes fixation and stereologic examination have already been detailed above. However, the studies by Sarkar et al. (1995, 1997) suffer from additional deficiencies that affect one's ability to draw conclusions from the data.

Although the authors note that the collected sperm were placed on a "warm slide" for assessment of motility, whether samples were randomized across groups for assessment is not known. Temperature can affect sperm motility (Peretz et al., 2012); thus, if the slides were not kept warm or were examined in group order, this may have affected the reported results. It also is not clear if motility was assessed using an automated system or manually. Although the stains used for assessing sperm morphology are appropriate, the percentages are different than typically recommended (Peretz et al., 2012) and the magnification at which morphology was assessed is not reported. Whether the samples were randomized or coded (to examine without reference to treatment) before examination also was not reported. Further, sperm morphology is typically a subjective assessment; thus, information regarding their classification scheme and how 'abnormal' was determined should have been included in the methods description.

With regard to the histoenzymology, the number of slides and sections evaluated per animal and the number of animals per group were not reported. Additionally, enzyme activity was evaluated based on “intensity of blue colour development”; however, how this was measured, whether it was done using an automated system, and if the activity was normalized to the activity of a housekeeping enzyme was not reported. Based on the numbering system used to report the findings in Table 2 of the study report, it would appear that enzyme activity was assessed subjectively, but methods to ensure an unbiased evaluation (i.e., blinded analyses) were not reported.

Finally, the number of animals per group is insufficient to draw any conclusions. Although 12 animals were included in each group, half of these animals were sacrificed after 7 weeks of treatment (Sarkar et al., 1995), leaving only 6 animals per group for analyses at 10 weeks of treatment. As noted in the footnote to Table 1 in Sarkar et al. (1997), the numbers provided in the table are mean values from 5-6 animals per group for sperm count, motility, abnormality and organ weight. For the other parameters, the means are derived based on only 3-5 measurements per group. These group numbers are far too low to take into account the potential variability in results.

Conclusions

As noted above, there are numerous fatal flaws in these experiments that preclude their use in safety assessment. The greatest of these involves (1) the apparent undernourishment of the rats used in the study, as evidenced by their low body weights and minimal weight gain over the course of the 10-week study; (2) the small number of animals/testes examined; and (3) inappropriate microscopy methods, including poor choice of fixatives and absence of record-keeping that would allow for the experiments to be reproduced. These failings make it impossible to understand and impartially assess the results. Consequently, the Sarkar et al. (1995, 1997) studies must be considered unreliable for assessing the potential male reproductive effects of isoproturon.

References

Brinkworth, MH, D Anderson, AE McLean. 1992. Effects of dietary imbalances in CD-1 mice and CD rats. Food and Chemical Toxicology 30:29-35.

Creasy, DM. 1997. Evaluation of testicular toxicity in safety evaluation studies: The appropriate use of spermatogenic staging. Toxicol Pathol. 25: 119-131.

Latendresse, JR, AR Warbrittion, H Jonassen, DM Creasy. 2002. Fixation of testes and eyes using a modified Davidson's fluid: Comparison with Bouin's fluid and conventional Davidson's fluid. Toxicol Pathol 30: 524-533.

Peretz, J, ZR Craig, S Bunting, JA Flaws. 2012. Experimental approaches to evaluate mechanisms of reproductive toxicity. Chapter 3, *In Developmental and Reproductive Toxicology. A Practical Approach.* 3rd Ed (RD Hood, Editor) Pp. 45-59.

Rehm, S, TE White, EA Zahalka, DJ Stanislaus, RW Boyce, PJ Weir. 2008. Effects of food restriction on testis and accessory sex glands in maturing rats. Toxicologic Pathology 36:687-694.

Sarkar, SN, AC Majumdar, SK Chattopadhyay. 1997. Effect of isoproturon on male reproductive system: clinical, histioenzyonological studies in rats. Indian Journal of Experimental Biology 35:133-138.

Sarkar, SN, SK Chattopadhyay, AC Majumdar. 1995. Subacute toxicity of urea herbicide, sioproturon, in male rats. Indian Journal of Experimental Biology 33:851-856.

U.S. Environmental Protection Agency (EPA). 2009. Pubertal development and thyroid function in intact juvenile/peripubertal male rats. OPPTS 890.1500. Endocrine Disruptor Screening Program Test Guidelines. EPA 740-C-09-012.

Weibel, ER, GS Kistler, WF Scherle. 1966. Practical stereological methods for morphometric cytology. J Cell Biol 30: 23-36.

Weibel, ER. 1969. Stereological principles for morphometry in electron microscopic cytology. Int Rev Cytol 26:235-302.

Weibel, ER and BP Bolender. 1973. Stereological techniques for electron microscopic morphometry. In Princples and Techniques of Electron Microscopy: Biological Applications, Vol 3, NA Hayat, Ed, Van Nostrand Rhinehold, NY, pp 237-296.