October 24, 2005

Guidance for Thyroid Assays in Pregnant Animals, Fetuses and Postnatal Animals, and Adult Animals

Office of Pesticide Programs
Health Effects Division
Washington DC

Purpose

The Food Quality Protection Act of 1996 (FQPA) provides enhanced protections for infants and children, directing EPA, in setting pesticide tolerances, to use an additional tenfold margin of safety to protect infants and children, taking into account the potential for pre- and postnatal toxicity and the completeness of the toxicology and exposure databases. The FQPA authorizes EPA to replace this tenfold "FQPA safety factor" with a different FQPA factor only if reliable data demonstrate that the resulting level of exposure would be safe for infants and children. Thus, an important consideration in implementation of the FQPA safety factor provision is assessing the completeness of the toxicity data with respect to prenatal and postnatal toxicity.

A number of pesticide chemicals have been shown to perturb thyroid hormone homeostasis via reduction of circulating thyroid hormones (Hurley *et al.,* 1998). Chemicals that perturb thyroid homeostasis and result in hypothyroidism are known to be associated with neurological disorders and alterations in neurological development, both in animals and humans (Fisher, 2000; Chan and Kilby, 2000; Morreale de Escobar *et al*., 2000; Zoeller and Rovet, 2004; Anderson *et al*., 2003). Thus, in the assessment of the toxic characteristics of a thyroid disrupting pesticide, determination of the potential to adversely impact thyroid hormones, thyroid structure, and/or thyroid hormone homeostasis during development is important. Normally, if a neurodevelopmental concern is raised by existing data on a pesticide, a rat developmental neurotoxicity (DNT) study is requested. However, disruption of thyroid homeostasis by thyroid disrupting pesticides is the initial, critical effect that may lead to adverse effects on the developing nervous system. Thus, *in lieu* of the rat DNT study, the special study described herein entails a mechanistic approach to generate specific data on the thyroid (*i.e.*, the primary target of the chemical of interest) to protect the developing nervous system from thyroid hormone disrupting chemicals. The specific purpose of this special study is to generate data to establish NOAELs and LOAELs (or benchmark doses) that may be used to derive RfDs that would be protective of the ability of a chemical to disrupt thyroid function in pregnant females and in the fetus and newborn. This special study will generally be requested based on the results of a study(ies) in adult animals that provide evidence that a pesticide produces effects on thyroid function or structure.

The test substance is administered to groups of pregnant animals during gestation and lactation. Unless specific data are provided that demonstrate sufficient exposure of the pups during lactation, a separate test group should be included that directly doses pups to ensure that there is sufficient exposure for a postnatal evaluation. The most reliable means to define postnatal exposure is to directly gavage the pup for some or all of early life (Chapin *et al*., 1997; Beyrouty *et al*., 2001; Moser *et al*., 2001). Thyroid evaluations are conducted on dams on gestation day (GD) 20 and postnatal day (PND) 21. Fetal thyroid evaluations are conducted at GD 20 in order to determine whether the test chemical alters fetal hormones or thyroid structure during this critical period of brain development (Bernal, 2002; Zoeller, 2003). Offspring are randomly selected from within litters for hormone evaluation at PND 4 and PND 21 to evaluate for effects on the thyroid and its function in early and late postnatal periods. The evaluation includes observations of circulating concentrations of triiodothyronine (T3), thyroxine (T4), and thyroid stimulating hormone (TSH), as well as histopathological assessments of the thyroid gland.

It should be stressed that this paper is considered guidance only and is not intended to be prescriptive. The Agency recognizes that registrants may chose to propose an alternative approach that adequately addresses the issue of disruption of thyroid hormone homeostasis during development. Registrants are encouraged to consult with the Agency during development of any alternative study protocol or the implementation fo the guidance provided here within.

Study Design Guidance

(1) Animal selection

(i) Species and strain. The most commonly used species for evaluating thyroid toxicity is the rat. If a sponsor wishes to use a mammalian species other than the rat, justification/reasoning for this selection should be provided.

(ii) Age. Young adult nulliparous females (approximately 3 months of age for rats) animals should be used.

(iii) Sex. For the main phase of the study, pregnant female animals are used.

(iv) Number of animals.

(A) The objective is for a sufficient number of pregnant animals to be exposed to the test substance to ensure that an adequate number of samples are produced for hormonal and histological evaluations. At least 40 litters are recommended at each dose level. 20 litters per treatment group should be assigned to the prenatal testing subset and 20 litters per treatment for the postnatal evaluations.

(B) On PND 4, the size of each litter should be adjusted by eliminating extra pups by random selection to yield, as nearly as possible, four males and four females per litter. Elimination of runts only is not appropriate. Whenever the number of pups of either sex prevents having four of each sex per litter, partial adjustment (for example, five males and three females) is permitted. Individual pups should be identified uniquely after standardization of litters (*e.g.*, Avery and Spyker, 1977).

(v) Assignment of animals for biochemical and pathological evaluations. Serum hormonal measurements (T4, T3, and TSH) and pathological evaluations of thyroid tissue should be performed on GD 20 dams and fetuses, on neonates on the day of standardization of the litters (PND 4), and on dams and pups on PND 21, for a total of 10 litters per treatment on GD 20 and 20 litters per treatment on PND 4 and PND 21.

(vi) Animal care. Animal care and housing should be in accordance with the recommendations contained in the DHHS/PHS NIH Publication No. 86–23, 1985, Guidelines for the Care and Housing of Laboratory Animals, or other appropriate guidelines.

(2) Control group

(i) A concurrent control group is required. Animals in the control group should be handled in an identical manner to test group animals.

(ii) If a vehicle is used to facilitate dosing, consideration should be given to the following characteristics: effects on the absorption, distribution, metabolism, or retention of the test substance; effects on the chemical properties of the test substance which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals. The vehicle should not be developmentally toxic, have effects on reproduction, nor affect the structure or function of the thyroid gland.

(iii) The vehicle control group should receive the vehicle in the highest volume used.

(3) Dose levels and dose selection

(i) At least three dose levels of the test substance plus a control group (vehicle control, if a vehicle is used) should be used. The dose levels should be spaced to produce a gradation of toxic effects. If available, information on metabolism and pharmacokinetics may be useful in the selection of doses.

(ii) Highest dose level. Because it is anticipated that thyroid data will be available from adult animal studies, a dose that produced a robust, toxicologically significant effect on thyroid hormone levels in the adult animal is the recommended highest dose level. This dose should not produce severe maternal or fetal toxicity. Note that the normal serum thyroxine concentrations in pregnant rats between GD 17-22 are significantly lower than in non-pregnant female rats (Fukada *et al.*,1980; Calvo *et al*., 1990; Versloot *et al*., 1994). This is in contrast to pregnant women, in which serum T_3 and T_4 rise progressively during gestation due to a thyroid-stimulating action of placental chorionic gonadotrophin (which is absent in pregnant rats).

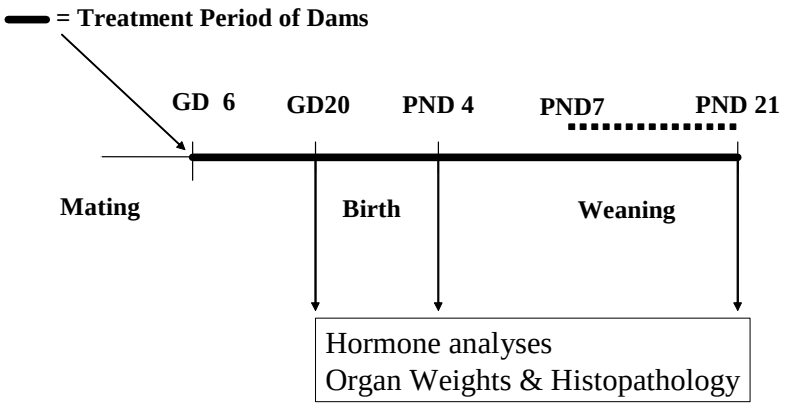
(iii) The lowest dose should not have any impact on thyroid hormones or thyroid structure.

(iv) The intermediate dose should produce detectable perturbation of thyroid hormone levels in the adult animal, in order to define the dose-response.

(4) Dosing period (Figure 1). Day 0 of gestation is the day on which a vaginal plug and/or sperm are observed. The maternal dosing period should cover the period from day 6 of gestation through day 21 postnatally. Dosing should be avoided, when possible, on the day of parturition.

If direct dosing of the pups via gavage is needed, it should start around PND 7 and continue to PND 21 (*i.e.*, a total of 14 days of dosing).

Figure 1. Developmental Thyroid Study



▪▪▪▪▪▪▪▪▪ = Direct dosing via gavage

Direct dosing is incorporated into the study to ensure adequate exposure to the offspring in the absence of sufficient lactation transfer. Care should be taken to minimize stress on the maternal animals and on the pups (Moser *et al.,* 2005).

(5) Administration of the test substance

(i) The test substance or vehicle should be administered by the principal route of potential human exposure; for pesticides, oral administration (gavage) is typical.

(ii) If another route of administration is used, the tester should provide justification and reasoning for its selection, and appropriate modifications to the study may be necessary. For materials administered by inhalation, whole-body exposure is preferable to nose-only exposure due to the stress of restraint required for nose-only exposure. The same route of exposure should be used for adults and offspring.

(iii) The test substance should be administered at approximately the same time each day.

(iv) When administered by gavage or dermal application, the dose to each animal should be based on the most recent individual body weight determination.

(6) Observation, termination, and parameters to be measured in test animals

(i) Maternal

(A) Each animal should be observed at least once daily, considering the peak period of anticipated effects after dosing. Mortality, moribundity, pertinent behavioral changes, and all signs of overt toxicity should be recorded at this cageside observation. Any signs of toxicity should be recorded as they are observed, including the time of onset, degree, and duration.

(B) Animals should be weighed on day 0, at termination, and at least at 3–day intervals during the dosing period.

(C) Food consumption should be recorded on at least 3-day intervals, preferably on days when body weights are recorded.

(D) Termination schedule.

(1) At least 10 pregnant dams per dose and control group should be terminated on GD20, immediately prior to the expected day of delivery. An additional 20 dams per dose and control group should be terminated on PND 21.

(2) Dams showing signs of abortion or premature delivery prior to scheduled termination should be sacrificed and subjected to a thorough macroscopic examination.

(E) Gross necropsy. Dams sacrificed at the scheduled termination or dying during the study should be examined macroscopically for any structural abnormalities or pathological changes that may have influenced the pregnancy.

(F) At each sacrifice time, maternal blood should be collected for biochemical analyses.

(G) At each sacrifice time, maternal thyroids should be collected for pathological analyses.

(ii) Fetal (GD 20)

(A) The sex and body weight of each fetus should be determined.

(B) Each fetus should be examined for external anomalies.

(C) Fetal blood should be collected and pooled by sex within litters for biochemical analyses.

(D) Fetal thyroids should be collected from one randomly selected male and female per dam for pathological analyses.

(iii) F1 Offspring

(A) All offspring should be examined cage-side at least daily for signs of mortality or morbidity. Any signs of toxicity in the offspring should be recorded as they are observed, including the time of onset, degree, and duration.

(B) Offspring should be weighed on the day of litter standardization (PND 4) and at least twice a week until PND 21.

(C) On PND 4 and PND 21, pup blood should be collected from one randomly chosen male and female offspring per litter. If necessary to increase sample volume, blood from all culled pups may be pooled by sex within litters.

(D) Pup thyroids should be collected on PND 4 and on PND 21 from one randomly selected male and female per litter for pathological analyses.

(iv) Animal sacrifice. Common methods of animal sacrifice include decapitation or sacrifice using anesthesia. If animals are anesthetized, information should be provided to show that the anesthetic does not impact thyroid hormones (Dohler *et al*., 1979).

***(v) Hormone analyses*.** Due to the circadian rhythm of thyroid hormones, sample collection should occur at approximately the same time of day and be randomized across dosage groups, preferably in the morning hours at which time basal values should be present (Dohler *et al*., 1979). Blood samples for evaluation of triiodothyronine (T3), thyroxine (T4), and TSH should be collected immediately following sacrifice. Hormonal analyses should be conducted on GD 20 fetuses and dams, PND 4 pups, and PND 21 pups and dams.

If there are inadequate fetuses in a litter to obtain sufficient blood for the hormonal measures, the measurements of T4 and TSH (if feasible) would be a priority, with less emphasis placed on T3 measures. If necessary, PND 4 pups of both sexes may be pooled to achieve sufficient blood. Prior to sacrifice, every effort should be made to avoid inducing stress that could affect hormone concentrations (Dohler *et al*., 1979).

(vii) Organ Weights and Histopathology. Thyroid and liver organ weights should be collected, as well as body weights. The thyroids/parathyroids, attached to a section of the trachea, of all maternal, fetal, and offspring samples should be excised and immersion fixed immediately after collection in 10% neutral buffered formalin or other appropriate fixative. Following fixation, the thyroid/parathyroid tissue samples should be carefully trimmed and weighed. Fixed thyroid tissue should be processed with routine embedding (*e.g.*, paraffin), sectioned appropriately (*e.g.*, 5 micrometer), and stained with hematoxylin and eosin (or other appropriate stains) and examined by light microscopy. The frequency of each type and the severity of each thyroid follicular alteration should be recorded. The spectrum of alterations should include colloid depletion, follicular cell size including changes in cell height, and follicular cell hyperplasia. The tissue alterations should be recorded as incidence and, for hypertrophy and hyperplasia, also severity. One may also wish to include a severity score for the follicular lumen size which would be related to the amount of colloid present, or, in some cases, a morphometric analysis of follicular lumen area. In most cases, the rating scale consists of a numerical evaluation of 1+, 2+, and 3+, which relates to the degree of severity, ranging from very slight to very extensive (Hardisty and Boorman, 1990; Hooth *et al*., 2001). A proposed severity scoring system for rodent thyroid is: follicular cell hypertrophy is considered present when thyroid follicles are uniformly lined by tall cuboidal to columnar epithelium. The severity can be scored based on the numbers of follicles with the alterations: 1 - few follicles, 2 - approximately 50% of the follicles, or 3 - more than 75% of the follicles lined by hypertrophic epithelium. Hyperplasia can be graded as, 1 - scattered individual or sometimes adjacent follicles with focal hyperplasia within the follicle, 2 - a greater number of scattered individual affected follicles or foci of more than 2 hyperplastic follicles, or 3 - numerous hyperplastic follicles with focal microfollicular formation within the follicles.

(viii) To evaluate the overall health status of the pregnant animals/offspring, it would be desirable to include some additional standard developmental/reproductive toxicity parameters (*e.g.*, gross necropsy of organs, uterine weight, number of corpora lutea, number of implantations).

Data collection, reporting, and evaluation

Data should be reported individually and summarized in tabular form, showing for each test group the types of change and the number of animals displaying each type of change. The following specific information should be reported:

(1) Description of test system and test methods. A description of the general design of the experiment should be provided. This should include the following:

(i) A detailed description of the procedures used to standardize observations and procedures as well as operational definitions for scoring observations.

(ii) Positive control data from the laboratory performing the test that demonstrate the sensitivity of the procedures being used. These data will establish the competence of the test laboratory personnel in the evaluation of effects in fetal and neonatal animals. A suggested positive control is propylthiouracil, which is routinely used to disrupt thyroid hormones during development (*e.g.*, Goldey *et al*., 1995).

(iii) All available relevant historical control data from the laboratory performing the test. These data will establish document test norms for the testing laboratory.

(iii) Procedures for calibrating and ensuring the equivalence of devices and the balancing of treatment groups in testing procedures.

(iv) Any major deviations from this study design should be documented and explained.

(2) Results. The following information should be arranged by each treatment and control group:

(i) Tables containing data for each animal must be provided showing:

(A) Its identification number and the litter from which it came.
(B) Its body weight.
(C) Results of hormonal tests for each animal.
(D) A summary of the pathological finding in each animal.
(E) Time and cause of death (if appropriate); any neurological signs observed; a list of structures examined as well as the locations, nature, frequency, and extent of lesions.

(ii) Photomicrographs demonstrating typical examples of the type and extent of the pathological alterations observed.

(iii) Summary data for each treatment and control group must include:

(A) The number of animals at the start of the test.

(B) The body weight and food consumption of the dams during gestation and lactation.

(C) Litter size and mean weight at birth and throughout the lactation period.

(D) The number of animals showing each abnormal sign at each observation time.

(E) The percentage of animals showing each abnormal sign at each observation time.

(F) The mean number of subjects per group and standard deviation for each continuous endpoint at each observation time. These will include serum concentrations of triiodothyronine, thyroxine, and TSH, and thyroid weights (both absolute and relative).

(G) Results of hormonal tests should also include:

(1) a tabular or graphical representation of the standard curve with an indication of the measurable range of the assay and an explanation of the disposition of data from samples outside this range.

(2) if radioimmunoassay kits are used, the sensitivity at 90% binding, an estimate of the average intra-assay coefficient of variation, and if applicable the inter-assay coefficient of variation.

(3) a description of the arrangement of samples in each set of assays including the use of quality control samples to determine any intra-assay drift.

(H) Results of the pathological analyses should also include:

(1) The total number of animals in each treatment group in which any lesion was found, as well as the number of animals examined in that group.

(2) A tabular or graphical representation of the average number of animals in each treatment group affected by each different type of lesion, the location, frequency and average grade of each type of lesion.

(iv) Copy of the study and any amendments

Evaluation of data

(1) Evaluation of test results. The evaluation should include the relationship between the doses of the test substance and the presence or absence, incidence, and extent of any effects. The evaluation should include appropriate statistical analyses. The choice of analyses should consider tests appropriate to the experimental design and needed

adjustments for multiple comparisons. The evaluation should include the relationship, if any, between observed pathological and biochemical alterations.

(2) Use of Positive and Historical Control Data. When appropriate, positive control and historical control data should be used to enhance interpretation of study results. This data, when used, should be compiled, presented, and analyzed in an appropriate and relevant manner. In order to justify its use as an analytical tool, information such as the dates of study conduct, the strain and source of the animals, and the vehicle and route of administration should be included.

(3) Statistical analysis of the study findings should include sufficient information on the method of analysis, so that an independent reviewer/statistician can reevaluate and reconstruct the analysis.

(4) Absorption and bioavailability of the test substance in any study showing an absence of toxic effects.

References

Anderson GW, Schoonover CM, Jones SA. 2003. Control of thyroid hormone action in the developing rat brain. Thyroid 13:1039-1056.

Avery DL, Spyker JM. 1977. Foot tattoo of neonatal mice. Lab Anim Sci. 27:110-112.

Bernal J. 2002. Action of thyroid hormone in brain. J Endocrinol Invest 25:268–288.

Beyrouty P, Deschamps Y, O'Shaughnessy DO, Pedersen K-M. 2001. Effect of methyl parathion administration on cholinesterase activity in adult, juvenile, and neonatal rats. Neurotox Teratol 23:293.

Calvo R, Obregon MJ, Ruiz de Ona C, Ferreiro B, Escobar Del Rey E, Morreale de Escobar G. 1990. Thyroid hormone economy in pregnant rats near term: a "physiological" animal model of nonthyroidal illness? Endocrinology. 127:10-16.

Chan S and Kilby MD. 2000 Thyroid hormone and central nerevous system development. J Endocrinol 165:1-8

Chapin RE, Harris MW, Davis BJ, Ward SM, Wilson RE, Mauney MA, Lockhart AC, Smialowicz RJ, Moser VC, Burka LT, Collins BJ. 1997. The effects of perinatal/juvenile methoxychlor exposure on adult rat nervous, immune, and reproductive system function Fund Appl Toxicol 40, 138-157.

Dohler KD, Wong CC, von zur Muhlen A. 1979. The rat as model for the study of drug effects on thyroid function: consideration of methodological problems. Pharmacol Ther [B]. 5:305-318.

Fisher DA. 2000. The importance of early management in optimizing IQ in infants with congenital hypothyroidism. J Pediat 136:274-274.

Fukuda H, Ohshima K, Mori M, Kobayashi I, MA Greer. 1980. Sequential changes in the pituitary-thyroid axis during pregnancy and lactation in the rat. Endocrinology. 107: 1711-1716.

Goldey ES, Kehn LS, Rehnberg GL, Crofton KM. 1995. Effects of developmental hypothyroidism on auditory and motor function in the rat. Toxicol Appl Pharmacol. 135(1):67-76.

Hardisty JF, Boorman GA. 1990. Thyroid gland. In: *Pathology of the Fischer Rat Reference and Atlas*. Boorman GA, Eustis SL, Elwell MR, Montgomery CA, Jr., McKenzie WF (editors). Academic Press, San Diego CA. pp.520-525.

Hooth MJ, Deangelo AB, George MH, Gaillard ET, Travlos GS, Boorman GA, Wolf DC. 2001. Subchronic sodium chlorate exposure in drinking water results in a concentration-dependent increase in rat thyroid follicular cell hyperplasia. *Toxicol Pathol*.29: 250-259.

Hurley *et al*. 1998. Mode of carcinogenic action of pesticides inducing thyroid follicular cell tumors in rodents. *Environ. Health Perspect*. 106(8): 437-445.

Morreale de Escobar G, Obregón MJ and Sescobar del Rey F. 2000. Is neuropsychological development related to maternal hypothyroidism or to maternal hypothyroxinemia? J Clin Endocrinol Metab 85:3975-3987.

Moser VC, Barone Jr. S, Smialowicz RJ, Harris MW, Davis BJ, Overstreet D, Mauney M, and Chapin RE. 2001. The effects of perinatal tebuconazole exposure on adult neurological, immunological and reproductive function in rats. Toxicol Sci 62, 339-352.

Moser, VC, Walls, I, and Zoetis, T. 2005. Direct dosing of preweaning rodents in toxicity testing and research: Deliberations of an ILSI RSI expert working group. *Int J Toxicol* 24:87-94

Versloot PM, Gerritsen J, Boogerd L, Schr oder-van der Elst JP, Van Der Heide D. 1994. Thyroxine and 3,5,3'-triiodothyronine production, metabolism, and distribution in pregnant rat near term. Am J Physiol Endocrinol Metab. 267:860-867.

Zoeller, R.T. 2003. Transplacental thyroxine and fetal brain development. J Clin Invest 111:954–957.

Zoeller RT, Rovet J. 2004. Timing of thyroid hormone action in the developing brain: clinical observations and experimental findings. J Neuroendocrinol. 16:809-818.