Workshop Summary

Summary of the National Toxicology Program’s Report of the Endocrine Disruptors Low-Dose Peer Review

Ronald Melnick,1 George Lucier,1 Mary Wolfe,1 Roxanne Hall,1 George Stancel,2 Gail Prins,3 Michael Gallo,4 Kenneth Reuhl,5 Shuk-Mei Ho,6 Terry Brown,7 John Moore,8 Julian Leakey,9 Joseph Haseman,1 and Michael Kohn1

1National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA; 2University of Texas Health Science Center, Houston, Texas, USA; 3College of Medicine, University of Illinois, Chicago, Illinois, USA; 4University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, New Jersey, USA; 5College of Pharmacy, Rutgers University, Piscataway, New Jersey, USA; 6University of Massachusetts Medical School, Worcester, Massachusetts, USA; 7Johns Hopkins University School of Public Health, Baltimore, Maryland, USA; 8Sciences International, Inc., Alexandria, Virginia, USA; 9National Center for Toxicological Research, Jefferson, Arkansas, USA

At the request of the U.S. Environmental Protection Agency (U.S. EPA), the National Toxicology Program organized an independent and open peer review to evaluate the scientific evidence on low-dose effects and nonmonotonic dose–response relationships for endocrine-disrupting chemicals in mammalian species. For this peer review, “low-dose effects” referred to biologic changes that occur in the range of human exposures or at doses lower than those typically used in the standard testing paradigm of the U.S. EPA for evaluating reproductive and developmental toxicity. The demonstration that an effect is adverse was not required because in many cases the long-term health consequences of altered endocrine function during development have not been fully characterized. A unique aspect of this peer review was the willingness submission of individual animal data by principal investigators of primary research groups active in this field and the independent statistical reanalyses of selected parameters prior to the peer review meeting by a subgroup of statisticians. The expert peer-review panel (the panel) also considered mechanistic data that might influence the plausibility of low-dose effects and identified study design issues or other biologic factors that might account for differences in reported outcomes among studies. The panel found that low-dose effects, as defined for this review, have been demonstrated in laboratory animals exposed to certain endocrine-active agents. In some cases where low-dose effects have been reported, the findings have not been replicated. The shape of the dose–response curves for reported effects varied with the end point and dosing regimen and were low-dose linear, threshold-appearing, or nonmonotonic. The findings of the panel indicate that the current testing paradigm used for assessments of reproductive and developmental toxicity should be revisited to see whether changes are needed regarding dose selection, animal-model selection, age when animals are evaluated, and the end points being measured following exposure to endocrine-active agents.

Key words: androgen, antiandrogen, bisphenol A, developmental toxicity, endocrine disruptors, estrogen, in utero exposure, low-dose effects, multigeneration study, neonatal exposure, reproductive toxicity. Environ Health Perspect 110:427–431 (2002). [Online 12 March 2002] http://ehpnet1.niehs.nih.gov/docs/2002/110p427-431melnick/abstract.html

The peer-review panel (the panel) included individuals from academia, government, and industry with expertise in receptor/molecular biology, experimental and clinical endocrinology, reproductive and developmental toxicology, statistics, and mathematical modeling. The panel was divided into five subpanels: Bisphenol A; Other Environmental Estrogens and Estradiol; Androgens and Antiandrogens; Biological Factors and Study Design; and Statistics and Dose–Response Modeling. This peer review used a unique and novel approach to evaluate the validity of this very important and controversial environmental health issue. Fifteen principal investigators of primary research groups active in this field were asked by the organizing committee to provide their individual animal data on selected parameters for independent statistical reanalysis by the statistics subpanel prior to the meeting. Data were willingly submitted from 49 of the 59 selected studies. In general, certain requested data sets were not provided because the data were not available in an electronic format as specified by the statistics subpanel, or the raw data were in the possession of collaborators and could not be provided in the requested time frame. Studies for which requested data sets were not submitted by principal investigators for independent

Address correspondence to R. Melnick, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, North Carolina 27709 USA. Telephone: (919) 541-4142. Fax: (919) 541-3647. E-mail: melnickr@niehs.nih.gov

The Peer Review Organizing Committee included W. Allaben, National Center For Toxicological Research, Food and Drug Administration; C. De Rosa, Agency for Toxic Substances and Disease Registry; P. Frenner-Crisp, U.S. Environmental Protection Agency (currently at International Life Sciences Institute); L. Goldman, Johns Hopkins University; S. Inkster, U.S. Consumer Products Safety Commission; J. Kariya, U.S. Environmental Protection Agency; R. Kavlock, U.S. Environmental Protection Agency; G. Lucier, NIEHS (retired); R. Melnick, NIEHS (Chair); E. Muirong, Centers for Disease Control and Prevention; M. Wolfe, NIEHS; and R. Hall, NIEHS.

Received 28 June 2001; accepted 16 October 2001.
review by the statistics subpanel were used as background information by the panel. Besides submitting their raw data, principal investigators were asked to provide for each study responses to a list of 23 questions on issues relevant to the evaluation of endocrine low-dose studies. These questions addressed animal source and specification, animal husbandry, chemical characterization, administration of test agent, treatment of controls, evaluation of end points, and methods of data analysis. Investigators from these research groups were also available at the meeting to give formal presentations of their research groups were also available at the meeting to give formal presentations of their findings and to have informal discussions with individual subpanels. Because of the extreme rigor of this evaluation and the extensive analyses of raw data performed by the statistics subpanel, unpublished studies were also included in this peer review.

The selected studies included a) treatments with bisphenol A, diethylstilbestrol (DES), ethinyl estradiol, nonylphenol, octylphenol, genistein, methoxychlor, 17β-estradiol, and vinclozolin, or b) effects of diet or intrauterine position. Exposure periods included in utero, neonatal, pubertal, adult, in utero through neonatal, in utero through puberty, and in utero through adult. Requested parameters included organ weights (prostate, testis, epididymis, seminal vesicle, preputial gland, uterus, and ovary), perinatal measures (e.g., anogenital distance), pubertal measures (e.g., age at vaginal opening, first estrus, preputial separation, and testis descent), and other relevant factors (e.g., daily sperm production, sperm count, serum hormone levels, lymphocyte proliferation in response to anti-CD3, histopathology, estrous cyclicity, receptor binding, estrogen-receptor levels, gene expression, and volume of sexually dimorphic nuclei of the preoptic area of the hypothalamus). To conduct this evaluation within a reasonable time frame, this review focused on reproductive and developmental effects. The extensive literature on dioxin and dioxin-like compounds was excluded because the U.S. EPA was finalizing its extensive and rigorous reevaluation of dioxin risk. Phthalate esters were also excluded because separate evaluations on these compounds were being conducted by the NTP Center for the Evaluation of Risks to Human Reproduction. A future workshop may focus on low-dose effects of dioxin-like compounds.

The statistics subpanel analyzed the raw data from 39 of the 49 submitted studies over a 6-week period and provided results from these analyses to the other subpanels 1 week before the peer-review meeting. These analyses provide greater insight on the experimental data than is typically apparent in most peer-reviewed research articles; consequently, the statisticians’ report was critical for each of the subpanel reviews.

Prior to the meeting, the Dose–Response Modeling group provided theoretical dose–response models based on mechanisms of receptor-mediated processes, as well as empirical dose–response models of endocrine-related effects. Several important statistical issues were identified by the subpanel and are addressed in their report; these include study sensitivity (power), adjustment for litter effects, pooling of control groups, exclusion of statistical outliers, accounting for body weight differences on organ weight effects, appropriateness of the selected statistical methodology, and data heterogeneity across dose groups. All of these matters, plus experimental design and conduct issues, were considered by each of the subpanels in their evaluations of the individual studies during the peer review. The statisticians and modelers participated in the other subpanel reviews to ensure that their analyses and models were appropriately used by the subpanels. A manuscript on the approach and general findings of the statistics subpanel was published recently (2).

The panel evaluated data from the major selected studies that support the presence or absence of low-dose effects in laboratory animals and that would be relevant for human health assessments. Hard copies of the publications or reports of the selected studies were sent to each of the panel members 2 months before the peer-review meeting. Low-dose effects analyzed by the panel should be considered as effects occurring at no-observed-effect levels (NOELs) because this review did not distinguish adverse versus nonadverse effects. However, the panel did compare, when appropriate, its analyses to existing no-observed-adverse-effect levels (NOAELs) or lowest-observed-adverse-effect levels (LOAELs) reported by the U.S. EPA or others. The panel was also asked to consider biologic and mechanistic data that might influence the plausibility of low-dose effects and to identify study design issues or other biologic factors that might account for differences in study outcomes. Conclusions from the panel on the existence of low-dose effects and the shape of the dose–response curve for endocrine-active substances in the low-dose region were based on the totality of available knowledge.

This unique scientific peer review provided an extraordinarily rigorous, open, transparent, and objective evaluation of the scientific evidence showing the presence or absence of low-dose effects of endocrine-disrupting agents and an opportunity for participation by all stakeholders who had interest in this scientific review. The independently prepared reports of the subpanels, as well as information on the selected studies and requested parameters, can be accessed at the NTP web site (3). Alternatively, hard copies of the final report can be obtained by contacting the U.S. EPA Office of Prevention, Pesticides, and Toxic Substances docket—42288A, (202) 260-7099. Highlights of the findings of the subpanels are given below.

**Peer-Review Subpanel Findings**

**Bisphenol A.** On the basis of the U.S. EPA estimate that the LOAEL for oral exposure to bisphenol A in rats is 50 mg/kg/day, the subpanel used 5 mg/kg/day as a cutoff dose for low-dose effects, regardless of the route or duration of exposure or the age/life stage at which exposure occurred.

Several studies provide credible evidence for low-dose effects of bisphenol A. These include increased prostate weight in male mice at 6 months of age and advanced puberty in female mice after in utero exposure to 2 or 20 µg/kg/day, and low-dose effects on uterine growth and serum prolactin levels that occurred in F344 rats but not in Sprague-Dawley rats exposed to 0.5 mg/kg/day. The latter findings demonstrate a clear difference in sensitivity to the estrogenic effects of bisphenol A in these two strains of rats.

Several large studies in rats and mice, including multigenerational studies in Sprague-Dawley rats, found no evidence for a low-dose effect of bisphenol A, despite the considerable strength and statistical power that those studies represent. For those studies that included DES exposure groups, those that showed an effect with bisphenol A showed a similar low-dose effect with DES (e.g., prostate and uterus enlargement in mice); those that showed no effect with bisphenol A also found no effect with DES.

Discrepancies in experimental outcome among studies showing positive and negative effects of bisphenol A may have been due to different diets with differing background levels of phytoestrogens, differences in strains of animals used, differences in dosing regimen, and differences in housing of animals (singly vs. group). Although some studies attempted to replicate previous findings, body weights and prostate weights of controls differed between these studies. Studies also differed in the extent of analysis of dosing solutions.

The subpanel concluded that "there is credible evidence that low doses of BPA [bisphenol A] can cause effects on specific end points. However, due to the inability of other credible studies in several different laboratories to observe low dose effects of BPA, and the consistency of these negative studies, the subpanel is not persuaded that a
low dose effect of BPA has been conclusively established as a general or reproducible finding. Data are insufficient to establish the shape of the dose–response curve for bisphenol A in the low-dose region, and the mechanism and biologic relevance of reported low-dose effects are unclear. The subpanel identified areas for additional research that would clarify uncertainties about low-dose effects of bisphenol A. These areas include additional low-dose studies, including the development and use of sensitive and easily measured molecular end points, following in utero as well as early neonatal exposure to conclusively establish low-dose effects of bisphenol A as a general, reproducible phenomenon; pharmacokinetic data in multiple species and strains of animals to characterize fetal uptake, metabolism, and elimination of bisphenol A and its metabolites; mechanistic data on estrogen receptor occupancy during critical periods of development, effects of specific receptor antagonists, and responses in estrogen-receptor knockout mice; additional studies to ascertain the effects of endogenous hormone levels as a function of intrauterine position; studies to evaluate the effects of differing levels of estrogenic components present in various feeds; characterization of genetic and epigenetic factors that affect responses to bisphenol A and hormones in general—e.g., factors that lead to strain and species differences in sensitivity; mechanistic studies on the effects of bisphenol A on regulation of transcriptional activity from gestation through adulthood. Other environmental estrogens and estradiol. The subpanel developed an operational definition for low-dose effects that was based on the dose–response data for the selected end points for each agent under evaluation. Low-dose effects were considered to be occurring when a nonmonotonic dose response resulted in significant effects below the presumed NOEL expected by the traditional testing paradigm. Low-dose effects were clearly demonstrated for estradiol and several other estrogenic compounds. The shape of the dose–response curves for effects of estrogenic compounds varies with the end point and the dosing regimen. Theoretical models based on mechanisms of receptor-mediated processes, as well as empirical models of endocrine-related effects, produced dose–response shapes that were either low-dose linear, or threshold-appearing, or nonmonotonic (e.g., U-shaped or inverted U-shaped). Low-dose effects of the estrogenic agents evaluated by the subpanel include the following: For estradiol (ovarian steroid with greatest estrogenic activity), low-dose effects include changes in serum prolactin, luteinizing hormone, and follicle-stimulating hormone in ovariectomized rats at a dose of approximately 3 μg/kg/day. For DES, a nonsteroidal synthetic estrogen that had been used to prevent spontaneous abortions and to enhance cattle weight gain, is a transplacental carcinogen in humans. There is clear evidence of a low-dose effect on prostate size after in utero exposure of mice to DES at 0.02 μg/kg/day. For genistein (isoflavone derived from soy), low-dose effects were observed in F1 offspring following dietary exposure (in utero through puberty) to 25 ppm. These effects include a decrease in the volume of sexually dimorphic nuclei of the preoptic area (SDN-POA) of the hypothalamus in male rats (approaching femalelike volumes), changes in mammary gland tissue in male rats, and an increase in proliferation of splenic T-lymphocytes stimulated with anti-CD3. For methoxychlor (insecticide), classic estrogenic activity occurs in F1 rats following in utero and perinatal exposure to 5 mg/kg/day or higher doses. Low-dose immune system effects occur in F1 offspring following dietary exposure (in utero through puberty) to 10 ppm methoxychlor (approximately equal to 1 mg/kg/day). For nonylphenol (industrial compound identified in drinking water supplies), low-dose effects in F1 rats following dietary exposure (in utero through puberty) to 25 ppm include a decrease in SDN-POA in males, an increase in relative thymus weight, an increase in proliferation of splenic T-lymphocytes stimulated with anti-CD3, and a prolonged estrus in females. For octylphenol (an intermediate for the production of surfactants), there was no evidence of low-dose effects in a five-dose multigeneration study in rats. Areas of future research include multiple dose studies and modeling of dose–response relationships; need for replication of low-dose findings in other studies or in other laboratories; determination of the toxicologic significance of volume changes in SDN-POA in male rats and the relationship between estrogenic activity and stimulation of lymphocyte proliferation. Androgens and antiandrogens. The subpanel’s review focused on low-dose effects of vinclozolin, a fungicide that is an androgen receptor antagonist. NOAELs for vinclozolin were established from studies in rats; these levels are 6 mg/kg/day for acute dietary exposure and 1.2 mg/kg/day from chronic dietary exposure. No studies have been conducted on vinclozolin at doses below its NOAEL. Exposure of pregnant rats to vinclozolin at six doses ranging from 3.125 to 100 mg/kg/day results in reduced anogenital distance (femalelike), increased incidences of areolas and nipple retention, and permanently reduced ventral prostate weight in male offspring. For these effects, the dose–response curves appeared linear to the lowest dose tested. Reproductive tract malformations and reduced ejaculated sperm numbers were observed only at the two highest doses. Thus, dose–response relationships are not equivalent among end points affected by exposure to vinclozolin. Antiandrogens act as androgen receptor antagonists, inhibitors of 5α-reductase activity, and/or inhibitors of steroidogenesis. In addition to vinclozolin, other agents (or their metabolites) that have been identified as antiandrogens include p,p’-dichlorodiphenyltrichloroethane (insecticide), flutamide, and Casodex (pharmaceuticals developed to treat prostate cancer), finasteride (pharmaceutical developed to treat benign prostate hyperplasia), methoxychlor (pesticide), procyomidine (fungicide), linuron (herbicide), ketoconazole (fungicide), and certain phthalate esters (plasticizers). For finasteride, which acts as a 5α-reductase inhibitor, the dose response for reduction in anogenital distance (linear) was different than that for increased hypospadias (threshold-appearing). No data are available on low-dose effects of environmental chemicals that act as androgen mimics. Future research needs include the following: further testing of the hypothesis that the dose response for antiandrogens is linear to the NOAEL/LOAEL; development of mechanism-based assays for the detection of androgen mimics; development and use of molecular and biochemical markers as sensitive indicators of low-dose effects of androgenic and antiandrogenic agents; characterization of dose–response relationships for androgenic and antiandrogenic agents in different species and in multiple strains; development of dosimetry/mechanistic models for exposures occurring during in utero and early neonatal development. Biological factors and study design. Several factors may account for discrepant findings on low-dose effects of particular endocrine-active agents. These factors include intrauterine position, which (although not essential for the detection of low-dose
effects) may be important in evaluating variability in response because differences in fetal exposure to endogenous hormones may influence responses associated with exposure to endocrine-disrupting chemicals;
• strain and strain differences in response, which could occur because of genetic differences or selective breeding to maintain high rates of fecundity and growth;
• diet with varying background levels of phytoestrogens and differences in caloric intake, which might influence reproductive parameters;
• differences in caging (e.g., stainless steel, polycarbonate), bedding material, or housing (group versus individual), which could influence study outcomes;
• seasonal variation, which has been reported to affect sex ratios in rodents.

Comments on the multigeneration test.
The traditional multigeneration reproduction study protocol includes exposure of animals through most critical windows of sexual differentiation in the F1 generation and an assessment of the F2 generation through postnatal day 21. This protocol provides substantial information on reproductive effects, but limited information on developmental effects. Frequently, litter size is reduced on postnatal day 4 (usually to four males and four females), and litter size is further reduced at weaning (postnatal day 21), so that only one animal/sex/litter is held until adulthood. The reduction in number of treated animals evaluated may provide inadequate power to detect low-incidence responses (e.g., reproductive tract malformations). Further, a number of sensitive or subtle endocrine-related end points are not routinely evaluated, and evaluations of F2 pups on or around postnatal day 21 may not reveal effects on reproductive tract organs that are not yet fully developed. This concern is underscored by the fact that certain endocrine-active chemicals were negative in standard multigeneration and prenatal studies.

Additional design factors for future studies include the following:
• Because of clear species and strain differences in sensitivity, animal-model selection should be based on responsiveness to endocrine-active agents of concern (i.e., responsive to positive controls), not on convenience and familiarity.
• Pharmacokinetic data need to be routinely generated, using appropriately sensitive methods, to characterize the dosimetry of the test chemical or its metabolites in target tissues.
• Caution is needed in implementing experimental designs to reduce animal variability (e.g., controlled feeding, individual housing), because factors such as body weight and stress can influence reproductive end points.
• The biologic/toxicologic relevance of specific end points affected by endocrine-active agents would benefit from measuring functional parameters or collecting mechanistic data on related biomarkers of effect.
• The long-term health consequences of early changes induced by endocrine-active agents, e.g., prostate enlargement or accelerated uterine development, need to be determined.
• Windows of susceptibility to endocrine-disrupting chemicals need to be identified from mechanistic data and empirical tests need to include exposures at those times.

Overall Conclusions
Low-dose effects, as defined for this review, were demonstrated in laboratory animals exposed to certain endocrine-active agents. The effects are dependent on the compound studied and the end point measured. In some cases where low-dose effects have been reported, the findings have not been replicated. The toxicologic significance of many of these effects has not been determined. The shape of the dose–response curves for these effects varies with the end point and dosing regimen, and may be low-dose linear, threshold-appearing, or nonmonotonic.

The traditional multigeneration reproduction study protocol has not revealed major reproductive or developmental effects in laboratory animals exposed to endocrine-active agents at doses approaching their NOAELs set by the standard testing paradigm. However, few multigenerational studies have been conducted over expanded dose ranges, and end points such as cancer of reproductive organs or neurobehavioral effects are generally not evaluated in multigenerational studies.

Subpanels.

Biphenol A
George Stancel (Chair) University of Texas at Houston
Gail Prins (Rapporteur) University of Illinois at Chicago
Nancy Cooper U.S. Environmental Protection Agency
Warren Foster Health Canada
Jan Kanno National Institute of Environmental Health Sciences
John Faust California Environmental Protection Agency

Other Environmental Estrogens and Estradiol
Michael Gallo (Chair) UMDNJ-Robert Wood Johnson Medical School
Kenneth Reuhl (Rapporteur) Rutgers University
Mari Golub California Environmental Protection Agency
Alaide Hughes UCLA School of Medicine
Richard Lyttle Wyeth-Ayerst Research
Lynne McGrath Schering-Plough Research Institute
Patricia Whitten Emory University

Androgens and Antiandrogens
Shuk-Mei Ho (Chair) University of Massachusetts Medical School
Terry Brown (Rapporteur) Johns Hopkins University School of Public Health
George Daston The Procter & Gamble Company
Mitch Eddy National Institute of Environmental Health Sciences
Lorentz Rhomberg Gradient Corporation
Elizabeth Wilson University of North Carolina at Chapel Hill

Biological Factors and Study Design
John Moore (Chair) Sciences International, Inc.
Julian Leakey (Rapporteur) National Center for Toxicological Research
Sue Barlow Consultant
Paul Foster Chemical Industry Institute of Toxicology
Robert Laueke U.S. Environmental Protection Agency
Robert Murphy National Institute of Environmental Health Sciences
Cory Teuscher University of Illinois at Urbana-Champaign

Statistics and Dose–Response Modeling
Joseph Haseman (Co-chair, Statistics) National Institute of Environmental Health Sciences
John Baier Miami University of Ohio
Ralph Kedell National Center for Toxicological Research
Richard Morris Analytical Sciences, Inc.
Kenneth Portier University of Florida
Michael Kohn (Co-chair, Modeling) National Institute of Environmental Health Sciences
Hugh Barton U.S. Environmental Protection Agency
Jim Coglanio U.S. Environmental Protection Agency
Rory Connolly Chemical Industry Institute of Toxicology
Robert Delongchamp National Center for Toxicological Research

430
The panel recommended additional research to replicate previously reported key low-dose findings, to characterize target tissue dosimetry during critical periods of development, to identify sensitive molecular markers that would be useful in understanding mechanistic events associated with low-dose effects, and to determine the long-term health consequences of low-dose effects of endocrine-active agents.

The findings of the panel indicate that the current testing paradigm used for assessments of reproductive and developmental toxicity should be revisited to see if changes are needed regarding dose selection, animal model selection, age when animals are evaluated, and the end points being measured following exposure to endocrine-active agents.

REFERENCES AND NOTES
1. U.S. EPA. Health Effects Test Guidelines OPPTS 870.3800 Reproduction and Fertility Effects. EPA 712-C-98-208. Washington, DC: U.S. Environmental Protection Agency, 1998.
2. Haseman JK, Bailer AJ, Kodell RL, Morris R, Portier K. Statistical issues in the analysis of low-dose endocrine disruptor data. Toxicol Sci 61:201–210 (2001).
3. National Toxicology Program homepage. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available: http://ntp.niehs.nih.gov/ [cited 1 August 2001].