
Reproductive Assessment by Continuous Breeding: Evolving Study Design and Summaries of Ninety Studies

Author(s): Robert E. Chapin and Richard A. Sloane

Source: *Environmental Health Perspectives*, Vol. 105, Supplement 1 (Feb., 1997), pp. 199-205

Published by: The National Institute of Environmental Health Sciences (NIEHS)

Stable URL: <http://www.jstor.org/stable/3433407>

Accessed: 11/11/2008 12:23

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=niehs>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



The National Institute of Environmental Health Sciences (NIEHS) is collaborating with JSTOR to digitize, preserve and extend access to *Environmental Health Perspectives*.

Reproductive Assessment by Continuous Breeding: Evolving Study Design and Summaries of Ninety Studies

Robert E. Chapin and Richard A. Sloane

National Institute of Environmental Health Sciences and National Toxicology Program, Research Triangle Park, North Carolina

The Reproductive Assessment by Continuous Breeding (RACB) design has been used by the National Toxicology Program for approximately 15 years. This article details the evolutions in the thinking behind the design and the end points used in the identification of hazards to reproduction. Means of nominating chemicals are provided, and both early and current designs are described as well as some proposed changes for the future. This introduction is followed by a text and tabular summary of each study performed to date. We hope that this will not only be an explicit presentation of the findings of this testing program to date, but will help stimulate thinking about new ways to detect and measure reproductive toxicity in rodents, and help identify new relationships among the end points that are measured in such studies. — *Environ Health Perspect* 105(Suppl 1):199–395 (1997)

Key words: mice, dominant lethality, breeding studies, sperm measures, reproductive toxicity, estrous cycle, phthalates, fertility, glycol ethers, National Toxicology Program

Introduction

As part of its charge to test chemicals of concern for potential toxicity evaluates reproductive toxicity using the design Reproductive Assessment by Continuous Breeding (RACB). This two-generation study design was developed by the NTP for use in identifying potential hazards to toxic effects on male and/or female reproduction, to characterize that toxicity, and to define the dose–response relationships for each compound. These studies have been performed by laboratories under contract to the National Institute of Environmental Health Sciences (NIEHS) using Good Laboratory Practices.

RACB studies have been generating public sector data for approximately 15 years, and we felt that summaries of the results to date would be useful to the scientific community. Earlier reports have summarized the genesis of the design and some of the initial results (1,2). Additionally, the results of numerous individual RACB studies have appeared in the peer-reviewed scientific literature; each of these studies is referred to later in this paper.

Ninety studies are summarized here. Each study contains text and a tabular summary of the results for that individual study. By themselves, however, these summaries

Many people contributed to the successful conduct of these studies. The other project officers at NIEHS, James Lamb, Rick Morrissey, and Jerry Heindel, deserve particular credit for conceiving of these studies and their modifications, and assuring adherence to the protocol throughout. The two contract laboratories, Research Triangle Institute and Environmental Health Research and Testing, Inc., excelled in the preparation for, conduct of, and reporting on these studies. The principal investigators at these institutions and their senior staff are listed for each compound. At NIEHS, Brad Collins, Bill Jameson, Tom Goehl and Jean Bernheim provided excellent coordination of chemistry support. Doug Walters provided useful assessment of Health and Safety concerns for these contracts, and David Bridge and Monica Frost have proven outstanding in monitoring quality assurance activities. Jim Plautz made the first efforts to draw much of these data together, under the encouragement and support of Bern Schwetz. Finally, a special bow of thanks goes to Jerry Heindel for his constant encouragement and new thinking in all aspects of the work here, and to Mike Shelby for his encouragement and active support of these activities.

Address correspondence to Dr. R.E. Chapin, Reproductive Toxicology Group, National Institute of Environmental Health Sciences and National Toxicology Program, PO Box 12233, Research Triangle Park, NC 27709. Telephone (919) 541-3474. Fax (919) 541-4634. E-mail: chapin@niehs.nih.gov

Abbreviations used: RACB, Reproductive Assessment by Continuous Breeding; NTP, National Toxicology Program; NIOSH, National Institute for Occupational Safety and Health; NCTR, National Center for Toxicological Research; ATSDR, Agency for Toxic Substances and Disease Registry; CPSC, Consumer Product Safety Commission; U.S. EPA, U.S. Environmental Protection Agency; OSHA, Occupational Safety and Health Administration; NCI, National Cancer Institute; NLM, National Library of Medicine.

are incomplete. Thus, this introduction provides some context for these individual reports: it reviews the changing data needs since the inception of the RACB tests, chronicles some of the responsive evolutionary changes in the design, and provides some overview of the effects of some of the classes of compounds run through this design. This paper will not address the relationship(s) among the different end points; a complete evaluation of these relationships is being undertaken and will be reported separately (Chapin et al., unpublished data). Thus, the intent of this review is to alert the reader to the existence of these data, to summarize the data collected for each compound, and to provide some context for each study. Access information is also provided for those readers desiring further information on a particular chemical.

The first 48 studies were performed using mice due to their small size and lower cost (3). The subsequent realization that rats may more correctly identify human reproductive toxicants, and that regulatory agencies deal with rat data more frequently and easily have led to the increasing use of rats in RACB studies. All the studies use rats; almost all of the studies performed previously and those reported here used mice.

Before describing the components of an RACB study, let us briefly summarize key events in the conduct of a study, beginning with the selection of a chemical for study.

The RACB Test Process

Nomination

While the specifics of the selection process have varied from year to year, the public and other government agencies have always had the capability to nominate compounds for evaluation. Nominating and evaluating chemicals for testing was carried out primarily through the Chemical Evaluation Committee. This interagency committee was responsible for chemical nominations for most of the 1980s. In addition, during that time, reproductive toxicologists from various components of the NTP (consisting of NIEHS, the National Institute for Occupational Safety and Health [NIOSH], and the National Center for Toxicological Research [NCTR]) would meet approximately two to three times a year to review test results and discuss chemical nominations.

Currently, there are two methods by which chemicals are nominated. Nominations may be made to the Office of

Chemical Nomination and Selection (OCNS) in the Environmental Toxicology Program (NIEHS, ETP, PO Box 12233, Research Triangle Park, NC 27709).

It is preferred that these nominations indicate the number of people exposed to the compound, the commercial importance of the chemical (pounds produced, current uses), environmental occurrence, and a summary of current information about the toxicity of the chemical. Those chemicals nominated to the OCNS will be evaluated by the Interagency Committee for Chemical Evaluation and Coordination (ICCEC), composed of representatives from the Agency for Toxic Substances and Disease Registry, Consumer Product Safety Commission, Department of Defense, U.S. Environmental Protection Agency (U.S. EPA), NCTR, Occupational Safety and Health Administration, National Cancer Institute, NIEHS, NIOSH, and National Library of Medicine. This process is described more fully in the NTP Annual Plan (available from NIEHS, Central Data Management, A0-01, PO Box 12233, Research Triangle Park, NC 27709. Telephone: (919) 541-3419. Fax: (919) 541-3687. E-mail: cdm@niehs.nih.gov). Direct nominations from the public to the Reproductive Toxicology Group are also accepted.

Chemistry

Once selected, chemical procurement is handled by NTP chemists. The chemistry support contractor procures the compound, characterizes it, and performs initial formulation and stability studies. This contractor also provides formulation instructions for the test lab and analyzes selected dose formulations for the correct amount of the test article.

In Vivo Exposures

Exposure routes that have been used are feed, water, and gavage. Dermal exposures have not been used because when animals are cohoused (as they must be for RACB studies), oral ingestion is certain, which seriously confounds the interpretation of a dermal study. Because of the 30-week duration of most RACB studies, the inhalation route is generally considered prohibitively expensive.

For compounds that have few or no existing data, a short dose-range-finding (DRF) study is performed. Doses for the main study are selected based on these data and/or any existing literature. The main issue is setting the high dose. For

compounds with no pre-existing data or which are not expected to impact reproduction, the high dose is picked based on an expected difference in body weight; a 10% difference between the high dose animals and the controls is the target. If some reproductive toxicity is expected, then a high dose is selected in the expectation or hope of producing infertility by the end of the cohabitation period (infertility is defined as no live pups). Middle and low doses are chosen to be successive divisions by either two or three, depending on the anticipated slope of the dose-response curve.

The contract lab performs the study and provides information to the NTP project officer throughout the study. Decisions are made about whether to perform a cross-over mating, which dose groups to evaluate histologically, and which organs to evaluate using histology or other methods (i.e., immunohistochemistry, special sperm studies, etc.).

Reports

The contract lab provides a draft final report, which is reviewed by the project officer. The second draft of the report is sent to two independent reviewers who review it for scientific conduct, study interpretation, and conclusions. Their comments are incorporated and a final report is issued. Copies of this final report are retained by the NTP and are also sent to the National Technical Information Service (NTIS), which distributes a copy of any report for a fee to those who request it (NTIS, U.S. Dept. of Commerce, 5285 Port Royal Rd., Springfield, VA 22161. Telephone: 703 321-8547). Each study in this summary has an NTIS number that should be used when ordering from NTIS.

Additionally, to provide broader access to the data, selected studies have been and will continue to be published in the peer-reviewed scientific literature.

Study Design Evolutions

Examination of the chemical summaries show that different end points have been evaluated for different compounds. To some degree this is dependent on the design and data needs for each compound, but this also reflects the evolving uses of these data, and thus, the design of the study. Fewer data were collected in the early studies than in later studies.

A common terminology is used throughout this paper and in other discussions about RACB studies. A brief review of this terminology and a description of the

events in an RACB study would be helpful in interpreting the summaries that follow.

Each study is separated into four tasks, though not all tasks may be performed for a given compound:

TASK 1 is the dose-range-finding (DRF) portion of an RACB study. The end points for Task 1 are body weights and food and water consumption. In early studies, Task 1 was performed for 2 weeks and focused exclusively on body weights and food and water consumption for five to eight animals at each of five dose levels, and controls. Subsequently, it became clear that selected compounds were reproductive toxicants at exposure levels that produced no change in these end points. For such compounds, this kind of DRF data could lead (and did lead) to setting some or all dose levels so high that no pups were produced at all. For such compounds, it would be useful to have a preliminary evaluation of reproductive function. This led to the modified 4-week Task 1, consisting of a 1-week exposure followed by a 3-week cohabitation and exposure period, and birth of the pups. Thus, in addition to more data on weights and consumptions (which can change as the animals acclimate to the exposure), litter data at delivery can be used to set the high dose. This has proven quite useful for several compounds.

TASK 2 is the main portion of an RACB study. Mice that are 10 to 12 weeks old at the start of exposure are used as the first generation (F_0). In Task 2, control and three dose levels are used, with 20 male and 20 female rodents per dose level. In almost all the studies reported here, 40 control pairs were used for reasons given below. Exposure begins 1 week prior to cohabitation (to allow for any effects on ovulation or sperm motility to manifest), and then the animals are housed as breeding pairs for approximately 14 weeks. During this time of continuous chemical exposure, litters are produced approximately 3 to 4 weeks apart. Data collected on each litter include the study day of delivery, number of male pups, number of female pups, aggregate weight of each sex, and number of dead pups observed. Cannibalism of dead pups is recognized to contribute to a low proportion of dead pups being recorded; more interpretive attention is given to live pup number and weight. The pups are removed and humanely killed; the dam enters a postpartum estrous; and the pregnancy cycle begins anew. Normally, four to five litters are delivered per adult pair during the 14-week cohabitation period. Adult

body weights are taken after each litter (females, to avoid confounding effects of pregnancy) and at selected intervals throughout the study (males).

After 14 weeks, the pair is separated for 6 weeks, during which the female delivers and nurses to weaning any last litter she may have conceived just prior to the end of the cohabitation period. During this time, the litter and body weight data from Task 2 are summarized and sent to the NTP project officer (PO), who determines whether there has been a significant adverse effect on reproduction.

In the presence or absence of reproductive toxicity, the last litter is nursed by the dam and weaned at postnatal day 21. Pups are counted and weighed at intervals during the nursing period. Toxicities presenting during this period could represent late expression of gestational effects, could be due to lactational transfer of compound or active metabolite, or could reflect compromised milk quality. Primarily, data from the nursing period serve as a trigger for further investigations.

It had been noted that the number of pups per litter and the number of pairs delivering a litter both tended to decline with time, so that fewer pairs produced slightly smaller litters for litters four and five. Also, it was feared that in the presence of a reproductive toxicant, there would be insufficient animals to evaluate the second generation in the most affected groups. An alternative model was tried with rats: rearing the second litter for F₁ evaluation, rather than the fifth. It was found not to present any significant advantages and in rat studies, the last litter is routinely reared for second-generation evaluation.

TASK 3 is the crossover mating trial, performed to determine which sex has been affected by treatment (or which is more affected). This trial is performed after the last litter from Task 2 has been weaned at postnatal day 21. Generally, Task 3 has only been performed with a single exposed group (often, the high dose), and controls. Three groups are formed: control males × treated females, treated males × control females, and controls × controls. To obtain 20 pairs in each group, 40 control pairs are needed. Task 3 animals are cohabited for a week without being exposed to the test compound, and the females are subject to vaginal lavage each day, to check for sperm. The animals are separated when the female is sperm positive or after 1 week, whichever comes first. Thus, alterations in libido or mating success can be identified

in this task. The females are allowed to carry and deliver their litter, whereupon the pups are assessed as above and humanely killed. The F₀ animals can be killed and evaluated for histopathology at this point. In most of the studies reported here, this F₀ necropsy evaluation was not performed.

TASK 4 is the evaluation of the second generation. Exposure to the test compound starts at weaning, with each pup receiving the same exposure level as that given his or her parents. Body weights are collected at several times during the growth phase to adulthood. When the animals are approximately 74 (mice) or 80 (rats) days of age, they are cohabited within treatment groups (but avoiding sibling matings) for a week. As in Task 3, the females are subject to vaginal lavage daily, and the pair is separated when the female is sperm positive or at the end of 1 week. The female carries and delivers the litter, which is evaluated as above, and the pups are killed. Females are lavaged again after delivery

and resumption of cyclicity to assess the nature of the cycle (normal, altered). The adult F₁ animals are then killed and subject to necropsy. Histopathology is performed at the discretion of the PO.

First version. Early studies were intended primarily to identify hazards and took a somewhat minimalist approach. The intent was that an RACB study (Figure 1) would be the first study on a compound, not the last. That is, evidence of reproductive toxicity generated from this design would stimulate other studies to more fully characterize the effect, identify target sites, etc. Thus, Task 1 was 2 weeks long and collected data on food and water consumption and body weights. For Task 2, much of the focus was directed at functional effects. Thus, histopathology was rarely evaluated on F₀ animals at the end of Task 2 or Task 3, or was limited to controls and high dose animals if, indeed, it was evaluated. In the earliest studies, histopathologic evaluations were generally limited to

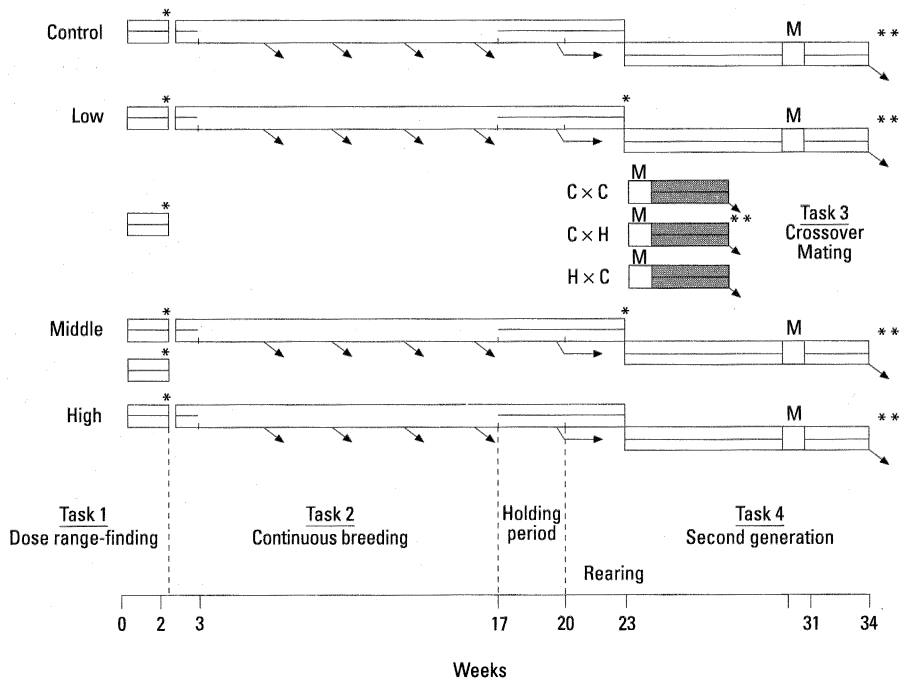


Figure 1. The original continuous breeding design. Task 1 is the 2-week range-finding segment. Task 2 begins with a 1-week dosing while the animals are housed separately (represented by the small horizontal line dividing each bar between weeks 2 and 3), whereafter one male and one female are housed as a breeding pair for 14 weeks under continuous exposure to the test chemical. The animals are separated again after 14 weeks of exposure and kept for a 3-week holding period, followed by 3 weeks to allow for the rearing of the last litter. The second generation begins when the animals are weaned (~ study week 23) and begin exposure to the same levels as those received by their parents. In Task 4, there is a single, 1-week mating trial, followed by separation until birth and evaluation of the litter. Task 3 would cross-mate treated animals of one sex with control animals of the other (see text for more complete description). Angled descending arrows indicate the birth of a litter of pups. M, mating period. *, the animals are killed and discarded; **, the animals are killed and a necropsy is conducted from Morrissey et al. (3).

controls and high dose animals at the end of Task 4. In some studies (~1985–1988), limited necropsy data were collected from all dose groups in Task 4.

Differences in responses between generations was not considered a likely event, therefore, identifying those differences was not a high priority. Thus, if a study found no effects on reproduction during Task 2 (that is, if Task 2 was negative), Task 4 would use only the control and high dose groups. This was a logical cost-containment strategy: if a trans-generational difference was unlikely and no effects were seen at any dose in Task 2, labor and money could be saved by not dosing and maintaining two groups of Task 4 animals that would likely not be affected by treatment. Differences in response could still be compared using the high dose group. If toxicity was observed during Task 2, all dosed groups would be evaluated in Task 4, though post-mortem evaluations might be limited.

In Task 3, there was a need for 40 control animals of each sex (20 to mate with a treated partner, 20 to mate with a new control partner). These additional control pairs also provided additional statistical power and helped generate a large control database quickly in the early days of the design. Thus, early studies each used 40 control pairs during Task 2 to provide sufficient animals in the event that a Task 3 was needed. For all studies that did not involve Task 3, the extra 20 pairs of controls (aside from their statistical power contributions) were underutilized. In the late 1980s, it was decided to try purchasing young adult animals to act as controls in the event Task 3 was needed. This use of different-age mating pairs has proven successful: the number of pairs delivering a litter is equivalent in groups of same-age partners as in young-old pairs. Current studies use 20 control pairs for Task 2, and purchase additional controls as needed for Task 3.

Current version. The main effect of changes in design (Figure 2) involves the collection of data for more end points. Task 1 can now be a 4-week test, with a single mating trial to generate some fertility information. Since all current studies now use rats, the duration of Task 2 has been increased by 1 week, to accommodate the slightly longer gestation period. Necropsy data are collected on all groups at the end of both generations. For a positive study, the groups not involved in Task 3 are held with continued dosing, and a complete necropsy is performed on at least 10 animals per sex per dose level, with histopathology focusing

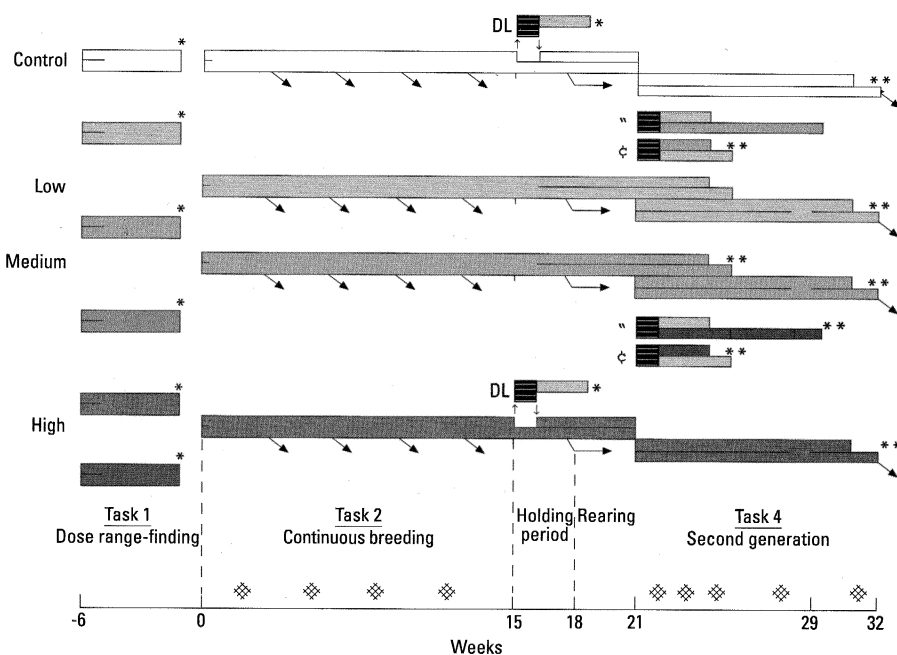


Figure 2. Current version of the reproductive assessment by continuous breeding design. See text for complete discussion of this design. The small horizontal line within each bar indicates separate housing; absence of a line indicates when the animals are cohoused as breeding pairs. Compared to the original version, Task 1 is longer, Task 2 is slightly longer to accommodate rat gestation length with the possibility of a dominant lethal segment (at the beginning of the holding period), and Task 3 is conducted using newly purchased, younger animals. The cross-hatched areas along the timeline indicate possible testing for screen grip strength. Descending angled arrows indicate birth of a litter. DL, dominant lethal; *, limited necropsy; **, full necropsy.

on reproductive and somatic target organs. This provides some dose-response data for end points that are thought to be more sensitive than rodent fertility, and 10 provides sufficient power to detect effects and estimate their prevalence.

It became clear during the course of these studies that functional changes in reproduction often were less sensitive than cell-based measures (sperm count, etc.). Thus, even if no functional changes are recorded during Task 2, there may be occult alterations in sperm indices or tissue structure. Thus, in a negative study (no adverse reproductive effects noted in Task 2), a limited necropsy is performed on 10 males in each dose group, taking sperm measures and reproductive organ weights.

In addition to the young-old pairing for Task 3, this crossover now also has the provision to further evaluate female reproduction. If implantation is hypothesized as a target, these animals could undergo a pseudopregnancy challenge test, to determine if there were treatment-related differences in the length of induced pseudopregnancy. This would provide a functional indication of altered hormonal status during

pregnancy. Alternatively, the females could be superovulated to assess their ability to ovulate after a hormonal stimulation. These two tests have yet to be successfully incorporated into an RACB study.

Finally, the NTP has long recognized that high quality histopathologic preparations can provide a great deal of information on the site of action of a toxicant. All testicular and epididymal tissues are routinely embedded and cut in glycol methacrylate and stained with periodic acid and Schiff's. This combination allows for the best possible routine evaluation of tissue structures. Additionally, the literature holds some examples of compounds that shorten reproductive lifespan by killing oocytes or otherwise depleting the ovary of oocytes. Counting and sizing follicles in serial sections of ovaries is another tool that can be used to determine site of effect.

Thus the end points for a current RACB study are shown in Table 1.

A change currently being considered is producing only three litters in the first generation, rearing the second generation from the third litter, and producing three litters in the second generation. This would

Table 1. End points for a current RACB study.

<p>Task 1 Feed and water consumptions Every 3–4 days Body weights Body weight gains Clinical signs</p>	<p>Task 3 necropsy (continued) Liver weight (absolute and relative) Kidney weight (absolute and relative) Weights of other target organs as appropriate Reproductive organ weights: ovary, testis, epididymis, seminal vesicles, prostate Testicular spermatid head count (expressed per mg tissue and per testis) Cauda epididymal sperm count Sperm motility (from vas deferens) by CASA Sperm morphology Histology as appropriate</p>
<p>Task 2 Feed and water consumptions during weeks of noncohabitation Pregnancy index by litter and dose Average number litters/pair Average number live pups/litter (male, female, combined) Proportion of pups born alive Sex ratio of liveborn pups Absolute live pup weight (by sex, and combined) Pup weight adjusted for litter size (by sex, and combined) Cumulative days to deliver each litter Adult body weights at delivery of each litter Feed and water consumption during lactation</p>	<p>Task 2 necropsy (performed in the absence of reproductive toxicity) Terminal body weight Gross observation of all organs and body cavities Reproductive organ weights: ovary, testis, epididymis, seminal vesicles, prostate Testicular spermatid head count (expressed per mg tissue and per testis) Cauda epididymal sperm count</p>
<p>Task 3 Mating index (% sperm positive females) Pregnancy index (% fertile pairs of those cohabited) Fertility index (number of fertile pairs per number sperm-positive females) Number live pups per litter Proportion of pups born alive Sex ratio of liveborn pups Absolute pup weight (by sex and combined) Average dam weight postpartum Average sire weight postpartum Average days to delivery</p>	<p>Task 4 F₁ pup body weights to weaning (pnd 1, 4, 7, 14, 21) F₁ pup viability to weaning (pnd 1, 4, 7, 14, 21) Mating index (% sperm positive females) Pregnancy index (% fertile pairs of those cohabited) Fertility index (number of fertile pairs/number sperm-positive females) Number live pups per litter Proportion of pups born alive Sex ratio of liveborn pups Absolute pup weight (by sex, and combined) Average dam weight postpartum Average sire weight postpartum Average days to delivery Feed and water consumption</p>
<p>Task 3 necropsy (performed in the presence of reproductive toxicity) Preterminal vaginal cytology for 12 days Terminal body weight Gross observation of all organs and body cavities</p>	<p>Task 4 necropsy as for Task 3 necropsy</p>

Abbreviations: CASA, computer-assisted sperm motility evaluation; pnd, postnatal day.

equalize the statistical power of both generations and would put more emphasis on functional effects after developmental exposure, a topic of significant current concern. The drawbacks of this approach would be that the second generation would not have been exposed from stem spermatogonia, but from committed spermatogonia. However, since very few compounds are stem-spermatogonia-specific toxicants, this would seem a small risk to run.

Integration with Other Tests

The RACB design generates three to four litters of young that are not kept for further evaluation. Additional developmental toxicity information can be gained from these studies through the use of one of these litters for structural evaluation of the pups. This biases the results because lethal alterations will be missed in this type of evaluation. However, lethal terata will manifest as reduced litter size, so the effect will still be

identified, even though a complete description will be lacking at this stage. Nonetheless, for those compounds that have no developmental toxicity data extant, the use of one of the litters for structural evaluation of all obtainable offspring offers the opportunity to glean at least screening-level information on the potential of the test compound to induce terata. Such a strategy is currently being pursued by the NTP.

The time between successive generations is sufficient to perform multiple additional evaluations of the animals on test. There are several effects that can be evaluated.

Neurotoxicity can be repeatedly assessed by a variety of measures (rotorod, grip strength, etc.), depending on the type of effect expected. These tests can be made at almost any point in the design, as they are noninvasive and repetitive (see studies on acrylamide and congeners).

When the F₀ mating pairs are separated at the end of Task 2, there is a 6-week

holding period during which the females are carrying and then nursing their young. During this time, the males are uninvolved. If there is prior suspicion that the test compound induces dominant lethal effects, new females can be purchased toward the end of Task 2, mated with these males, and killed before delivery to provide some measure of dominant lethality (DL). Alternatively, if no prior genetic toxicity data exist, a more logical sequence might be: perform Task 2, observe toxicity; perform Task 3, find male effects; then perform a dominant lethal test to test for DL in males.

In addition to generating data on untested compounds, the NTP is charged with developing new test methods. Two methods are being evaluated in collaboration with NIOSH. One of these is the sperm chromatin structure assay (SCSA) (5), which measures alterations in chromatin structure (relative abundance of single-stranded DNA vs double-stranded DNA). This test is being considered for inclusion in human field studies by NIOSH, but there is a relative paucity of data placing altered SCSA into some functional context. Because each RACB study develops extensive data on reproductive function, any changes in sperm SCSA could be compared to all the other data generated by the RACB design. Such a comparison would allow for an evaluation of the value added by use of SCSA in human field studies, as well as providing an indication of its benefit in rodent studies.

Another new method being evaluated by NIOSH for use with humans is sperm morphometry (measures of sperm head shape as opposed to shape classifications). Again, sperm from RACB animals are being used for morphometrics, and the additional data from the RACB study provide some context for these morphometric data.

Uses of RACB Data

Data from RACB studies form an effective part of the risk assessment process. These data identify hazards to reproduction, help characterize the toxic effects, and provide an indication of dose–response relationships. Data from these studies have been used in combination with other studies evaluated by the U.S. EPA and NIOSH to set acceptable exposure levels. These data also have provided the starting place for subsequent studies that have investigated the site and mechanism of a compound's toxicity.

Chemical Results by Class

Any testing program of this scope and with an open nominations process will evaluate

a wide variety of compounds for toxicity. Such is the case for the RACB program.

Not only were compounds evaluated individually for toxicity; several mixtures were assessed for their impact on reproductive and developmental processes. Additionally, the design was used to test the test species: a toxic glycol ether was used to evaluate the best design to use for rats, and three different strains of mice were evaluated to determine if a strain that was reproductively less robust might be more sensitive to compound-induced toxicity.

While most of these compounds were nominated individually, there are some class studies. Those compounds that were individually nominated and tested will not be reviewed here, as there is no common structural theme that links this group of miscellaneous compounds. However, the glycol ethers, phthalates, acrylamides, and mouse strain studies are four class studies that would benefit by a brief summarization of the effects overall.

Glycol Ethers

Ethylene glycol was found to produce facial abnormalities in offspring of treated mice, although the number of offspring was not reduced. Some ethers of ethylene glycol can be potent and effective reproductive toxicants. Those compounds with the shortest chain lengths are most toxic. Increasing chain length from monomethyl through monobutyl to monophenyl ethers decreased the degree of effects and increased the doses required to produce an effect on reproduction.

Diethylene glycol (DG) caused minimal reproductive toxicity at approximately 6 g/kg/day, while DG monoethyl ether caused no observable reproductive toxicity.

Propylene glycol (PG) had no adverse reproductive effect, while PG monomethyl ether caused a slight weight decrease in pups of treated dams at approximately 3 g/kg/day.

Triethylene glycol (TG) and TG diacetate were without effect, while TG dimethyl ether reduced fertility and pup number at 87 to 175 mg/kg/day.

Metabolites (methoxyacetic acid and ethoxyacetic acid) of active glycol ethers also impaired reproduction in ways quite similar to those seen with the parent molecule. It is clear that some of the short chain ethylene glycol ethers and their metabolites are reproductive and developmental toxicants in both males and females; the mechanism(s) of this toxicity is currently unknown. The absence of significant genotoxicity for this class (6) suggests a nongenomic interaction

that (based on the structures involved) is probably noncovalent. Additionally, there are clear structural determinants (longer side chains are less toxic), which suggests that a critical binding location (or more generically, a locus of interaction) does indeed exist. Changes in calcium flux appear to mediate some of the toxicity of ethylene glycol monomethyl ether (7), but this putative mechanism has not been investigated for any other glycol ethers to date.

Phthalates

Like glycol ethers, a number of phthalates were tested as a class of structures. These structures have a core benzyl ring with two identical substituent groups attached ortho to each other. To become active, however, one of these substituent groups is cleaved off at the ether linkage. The most toxic phthalates have 5- or 6-member side-chains (the di-N-hexyl and di-N-pentyl phthalates, respectively). Toxicity decreases with shorter chain lengths, suggesting (again) the presence of some structurally specific interaction with a target molecule. The nature of this molecule is still unknown.

Acrylamides

Acrylamide is both a neurotoxicant and an inducer of dominant lethal mutations in rodents. Based on data derived from relatively short-term exposures (8), the four studies summarized here were performed to explore structural correlates of these two toxicities, and to see if one effect could be produced in the absence of the other. All four studies employed the dominant lethal and grip strength evaluations mentioned above as additional evaluations during the in-life phase of the study. It was possible to separate the dominant lethality from neurotoxicity for this structural family: dominant lethality was seen in the absence of detectable neurotoxicity for methylene-bis-acrylamide, while neurotoxicity was detectable (to minimal degrees) with acrylamide and hydroxymethylacrylamide. Both hydroxymethylacrylamide and acrylamide itself produced significant dominant lethal effects, while methacrylamide was without measurable effects on reproduction in mice.

Mouse Strains

While most rodents have high fecundity, humans are thought to be reproductively less robust. These studies addressed the possibility that a less fecund strain should be the strain of choice for testing of chemical effects on reproduction. The question was: would strains of differing basal fecundity

respond differently to a toxicant? Three strains of mice (Swiss CD-1, C57Bl6, and C3H) were exposed to similar amounts of ethylene glycol monomethyl ether (EGME) in the drinking water. The most fertile strain (Swiss CD-1) was affected the least by EGME consumption, while the least fertile strain (C3H) showed greater reproductive toxicity to the same amounts of EGME. These studies are insufficient by themselves to fully assess the impact of using less fecund rodents routinely for testing. If the response to EGME is predictive of the response to other toxicants, one might predict that using less fecund strains would produce data of lower confidence (because of higher variability) and would probably alter the interspecies extrapolation factors, but would not likely improve the process of hazard detection.

Layout of the Summaries

Each compound presented here has a tabular summary of the observed effects. The format is designed to be intuitive to most readers: up arrows represent a significant increase, down arrows, a significant decrease. Solid dots indicate that no data were gathered for an end point in a dose group, while a horizontal dash indicates that no change was observed.

These tables present key information needed to understand the effects seen during the study, but not all end points are listed. If significant changes were seen in a nontabulated end point, they are addressed in the accompanying text summary of the study, which gives a rationale for each study, provides some quantitative idea of the magnitude of the changes that are dichotomized on the tables, and provides access information for each study in the header. Note also the dates for each study: early studies may have slightly different information than later studies.

Both text and tables mention only those effects where the treated group was statistically different from the controls at $p < 0.05$. There are a few instances (e.g., di-n-hexyl phthalate) where data from all groups are presented and only a few are significant. In these cases, the group that is different from controls has an asterisk indicating such.

Both the tables and the accompanying text refer to organ weights adjusted for body weight for all organs except testis. This approach is supported by data from several feed restriction studies that are also summarized below. These studies showed that reducing body weight gain by limiting feed availability and intake concomitantly

reduced organ weights. This was true for all organs examined except for testis, the weight of which remained constant until body weight gain was severely inhibited. Although additional insights might be gained by reporting both absolute and relative weights for all these organs, this presentation is meant to summarize the data, not report them exhaustively. The full data set is available for each compound through the sources mentioned earlier; those wishing to

compare absolute versus relative changes should consult the full report.

Similarly, pup weights could be expressed as either absolute pup weights or weights adjusted for litter size. Because litter size does affect the weight of each individual pup, we have chosen to mention only adjusted pup weights in these summaries. Although every effort has been made to make this explicit throughout the individual chemical reports, readers should

keep this in mind when reviewing the summaries. The hope is that this information will be useful in showing which compounds have been through such a testing scheme, in identifying which compounds cause what effects, and providing food for thought. Alert readers may identify trends that have been previously overlooked, and that this will stimulate new approaches and new ways of thinking about reproductive toxicity testing.

REFERENCES

1. Lamb JC IV. Reproductive toxicity testing: evaluating and developing new testing systems. *J Am Coll Toxicol* 4:163-171 (1985).
2. Reel JR, Lawton HD, Wolkowski-Tyl R, Davis GW, Lamb JCIV. Evaluation of a new reproductive toxicology protocol using diethylstilbestrol (DES) as a positive control. *J Am Coll Toxicol* 4:147-162 (1985).
3. Morrissey RE, Lamb JC IV, Morris RW, Chapin RE, Gulati DK, Heindel JJ. Results and evaluations of 48 continuous breeding reproduction studies conducted in mice. *Fundam Appl Toxicol* 13 :747-777 (1989).
4. Gulati DK, Hope E, Teague JL, Chapin RE. Reproductive toxicity assessment by Continuous Breeding in Sprague-Dawley rats: a comparison of two study designs. *Fundam Appl Toxicol* 17:270-279 (1991).
5. Evenson DP. Acridine orange-induced precipitation of mouse testicular sperm cell DNA reveals new patterns of chromatin structure. *Exp Cell Res* 187:328-334 (1990).
6. McGregor DB. Genotoxicity of glycol ethers. *Environ Health Perspect* 57:97-103 (1984).
7. Ghanayem BI, Chapin RE. Calcium channel blockers protect against ethylene glycol monomethyl ether-induced testicular toxicity. *Exp Mol Pharmacol* 52:279-290 (1990).
8. Hashimoto K, Sakamoto J, Tanii H. Neurotoxicity of acrylamide and related compounds and their effects on male gonads in mice. *Arch Toxicol* 47:179-189 (1981).