

# Identifying Environmental Factors Harmful to Reproduction

by Anthony K. Palmer

Reproduction is essential for the continuation of the species and for life itself. In biological terms, living and reproducing are essentially one and the same. There is, therefore, no sharp division between identifying factors harmful to reproduction and identifying factors harmful to life or vice versa. Detection of harmful factors requires balanced use of a variety of methodologies from databases on structure-activity relationships through *in vitro* and *in vivo* test systems of varying complexity to surveys of wildlife and human populations. Human surveys provide the only assured means of discriminating between real and imagined harmful factors, but they are time consuming and provide information after the harm has been done. Test systems with whole animals provide the best prospects for identifying harmful factors quickly, but currently available methods used for testing agrochemicals and drugs need a thorough overhaul before they can provide a role model. Whether there is a need for new methodology is doubtful. More certain is the need to use existing methodology more wisely. We need a better understanding of the environment—whatever it is—and a more thoughtful approach to investigation of multifactorial situations.

## Introduction

Reproduction ensures the continuity of the species, the race, the family, and of life itself. Living and reproducing are so interwoven that mammals, including humans, are structured physically, physiologically, and even psychologically to ensure efficient reproduction. The complexity of this interrelationship is such that separation of effects on reproduction from those affecting life in general is an artifice. Separation is a device of intellectual convenience that, if pursued too earnestly, can be misleading. It can be even more misleading to attempt to place different aspects of reproduction into distinct categories. For this reason, this paper may stray outside anticipated borders.

Concern about factors affecting reproduction is innate and inevitable. Concern about effects on reproduction, for which there is no immediate explanation, is even more inevitable. In trying to identify causes for unexplained effects, it is inevitable that society would look for any harmful factors in the environment. It is also inevitable that the objectivity of the search will be confounded by emotion and instinctive fears.

## Human Surveys

As to how we may detect and identify harmful factors, the first and most obvious method is direct observation of the reproductive outcomes in human populations. In earlier centuries these observations were often speculative and highly subjective. This often led to erroneous conclusions, such as attributing malformations to visitations from demons, or to nefarious sexual practices. Nowadays, attribution to the witch's curse would be unbelievable but, nevertheless, the witch hunt continues with chemicals and pharmaceuticals being the target. The continued legalistic attacks on the drug Bendectin provide a prime example of how primitive fears can override logic in even the most advanced societies (1-4).

Ultimately, for all its difficulties, direct observation of humans provides the only certain means of identifying factors harmful to human reproduction. Nevertheless, even with modern technology, it may take years rather than months to obtain certain results, and it still equates to "shutting the stable door after the horse has bolted."

## Wildlife Surveys

Another means of identifying environmental factors that might affect human reproduction is to monitor wildlife populations. Species at the end of food chains could be especially sensitive indicators of adverse environmental situations. Such species are perpetually close to the limit for survival, and the population density can be markedly affected by an effect on any of the species in the food chain.

---

Huntingdon Research Centre Ltd., P.O. Box 2, Huntingdon PE18 6ES, Cambridgeshire, UK.

This manuscript was presented at the Conference on the Impact of the Environment on Reproductive Health that was held 30 September-4 October 1991 in Copenhagen, Denmark.

An example often cited (5-7) is the effect of the persistent chlorinated hydrocarbon insecticides on populations of raptors, such as the kestrel, and on populations of small mammals and insectivores. Discontinuing the use of these insecticides has led to a marked recovery in populations of these species in countries such as Great Britain. Unfortunately, in some other parts of the world, discontinuing insecticide use also provided a reprieve for malaria carrying mosquitoes, with adverse consequences to humans. Also, replacement of DDT with more acutely toxic insecticides caused a number of deaths. Ironically, it is debatable whether human reproduction was ever affected by these insecticides.

Monitoring wildlife populations is notoriously difficult. It requires years of observation to avoid confusion with natural fluctuations in population density. It is available only for those countries that can afford it for the protection of wildlife itself. The most serious drawback would be the lack of understanding of both the vagaries of studying animal populations and of toxicology. At best, monitoring wildlife populations may provide a source of hypothesis generation regarding harmful factors in the environment.

## Testing with Model Systems

A third method of identifying adverse environmental factors is testing with model systems. The attraction is that testing offers the prospect of early identification, with greater economy, compared to surveys of human or animal populations. Testing offers the prospect of identifying adverse factors before human populations are seriously affected, not after. The confounding factors that bedevil the epidemiologist can be controlled to a considerable extent. The question is, can these prospects be realized in practice? This is the question I attempt to address here.

I would like to say that we have established an efficient test systems, but this would not be credible. The confounding factors of epidemiology are simply replaced by confounding factors of a different type. Identification of factors before humans are affected is feasible only for new factors. What about the many unexplored factors that already exist? For established environmental factors, we have to consider that the effects may be subtle and difficult to detect. Already, the more obviously harmful factors have been identified by direct observation of the human condition. These known factors would certainly include radiation, disease, pestilence, famine, malnutrition, ignorance, and poverty. Perhaps we should be a lot better at alleviating these more obvious harmful factors before adding further, more subtle ones.

We also have to ask the question of what test systems would be appropriate. I do not think that there has been any concerted effort to devise a test strategy specifically addressing the problems provided by the environment. What is the environment, anyway? The world provides us with an infinite variety of ever changing environments, but many people do not appear to be able to see beyond the factory wall, or even the laboratory bench. Whatever an individual's concept of the environment is, it is still a dynamic complex of interacting factors. Danger may lie

not in the presence of a factor but in the absence of one, for example, the absence of a trace element or nutrient. Danger may lie in an imbalance rather than the presence or absence of any one factor, and it may lie in the change rather than the situation at any one time.

Against this dynamic, multifactorial situation, most test systems are aimed at determining the effects of a single factor in a defined set of conditions. Perhaps more to the point, attitudes and philosophies have become conditioned to this mental "environment," and they may not be the most appropriate for testing environments.

In the absence of a specifically designed strategy for environmental factors, the most obvious role models are the regulatory tests required for chemicals and pharmaceuticals (8-11). In their favor, most, if not all, of the materials known to be harmful to human reproduction have shown adverse effects in these test systems. Most certainly, these testing methods have prevented the addition to the environment of some potentially harmful materials.

## Negative Aspects of Current Guidelines

Conversely, on the debit side, the test systems also have prevented the introduction of materials that would have been beneficial to humans. The problem is that there is no reliable measure of the proportion of harmful materials avoided to the proportion of beneficial materials lost. It seems a serious omission that we have no suitable mechanism to monitor the efficiency of these regulatory test methods and no suitable mechanism to control the idiosyncrasies of those who operate them. There is strong circumstantial evidence to indicate that, in the absence of a reliable reference point, testing methods for chemicals and pharmaceuticals have become inefficient, uneconomical, and, at times, ridiculous.

For example, for international registration of a new drug, it is not uncommon for more rats to be used for testing for effects on reproduction than for all other toxicity requirements combined. Manson (11) has quoted use of more than 6,000 rats as a general figure; numbers of 10,000 or more have been used in some cases. I know of at least one case where even more animals have been used as the manufacturer has striven to conduct guideline tests in rats when the rat was an unsuitable model for humans. Testing on this scale is unreasonable. Clearly, for testing all the unknowns in the environment it would be untenable.

In the absence of a reliable check system, guidelines also have accumulated a number of unvalidated and doubtful procedures (8,12,13) such as: a) the requirement for prolonged (9-10 weeks) premating of males, b) standardization or culling of litters, c) random selection of second-generation offspring, d) inclusion of interim sacrifices for fetal examination, and e) inadequate specifications for choice of dosages, etc. Explanation as to why these procedures can be considered as flaws is beyond the scope of this paper.

Also, in the absence of a check system, it would appear that many people have lost sight of the primary purpose of

testing, namely, to detect materials potentially hazardous to reproduction. The scientific content has been diminished, and, instead, testing has become an expensive game for administrators, lawyers, and officials. The fundamental similarities between the aims and scope of guidelines has been neglected. The small differences between them have been exaggerated and exploited. Quite clearly, if existing guidelines are to be used as a role model for testing environmental factors, it would be necessary to eliminate these negative features and get down to basics (8,12-16).

### Fundamentals

If we look at the basic similarities of these guidelines, they all attempt to cover all aspects of reproductive toxicity (Fig. 1). This is achieved either in one or in a combination of studies. Reproductive toxicity is often subdivided into two elements. One element concerns effects on the fully mature, functional adult, the other concerns effects on the developing organism. As mentioned earlier, this is a separation of convenience because the two elements are indivisible components of an integrated process.

For various reasons, most attention is given to effects that may be induced during the period of development from conception through puberty (Fig. 2). During this time the sensitivity and response to the environment may differ quite markedly from that of the mature, functional adult. It is important to note that development continues well after birth. Organic lead provides an example of a

substance for which exposure during infancy is most to be feared (17) and for which the consequences may not become apparent until later in life. Another example is diethylstilbestrol, for which exposure during perinatal stages of development results in delayed manifestations such as reproductive tract tumors of females appearing at puberty (18) and psychological disorders becoming apparent in early adult life (19).

The full scope of developmental toxicity, particularly the late manifestations of functional deficit, is not fully appreciated. For historical and psychological reasons, most attention is given to prenatal effects evident before or at birth. The greatest preoccupation is with localized effects on growth that result in birth defects. This concern with birth defects (teratogenesis) prompted the introduction of guidelines for reproductive toxicity and has greatly influenced their evolution ever since, although not necessarily for the better.

The emphasis on teratogenesis is completely out of proportion relative to the chance of occurrence. For many reasons teratogenesis is the least likely manifestation of developmental toxicity, and this is reflected in the low prevalence of coincidental malformations in all mammalian species (Fig. 3). Even magnitude increases in prevalence would have little impact on the viability of populations.

It is the rarity of malformation that causes devastating shock to the family. It is the shock that induces fear and dread that distorts rational assessment. The rarity of malformations makes for extreme difficulty in assured detection of an increased prevalence by direct observation. In fact, in testing, the indirect method of observing other manifestations that always associate with teratogenesis provides a more assured means of detection and discrimination. These other manifestations such as death, altered weight and increased prevalence of minor structural changes are also important in their own right. If efficient testing systems are to evolve, we need to arrive at a better balance between the real risk of teratogenesis and the exaggerated perception of risk that has prevailed to date.

In a broader perspective, to detect most of the direct and indirect ways that reproduction may be impaired, minimum requirements would include exposure from just

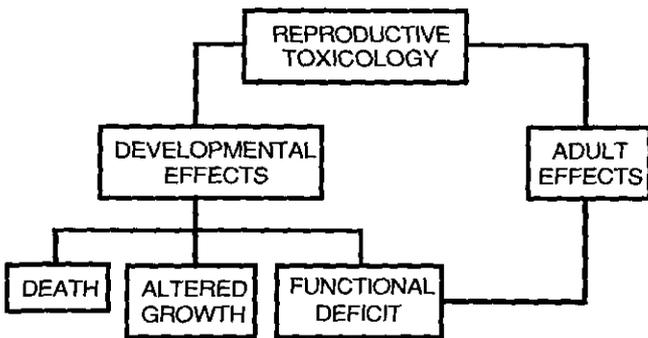


FIGURE 1. Major divisions of reproductive toxicity.

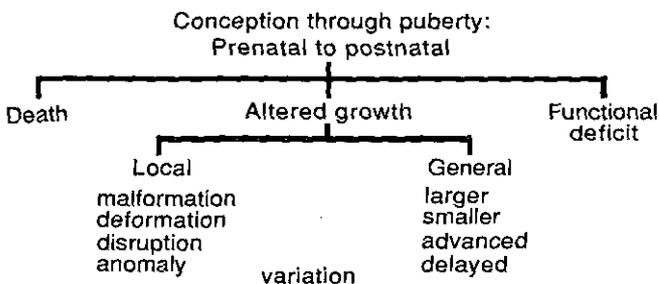


FIGURE 2. Manifestations of developmental toxicity.

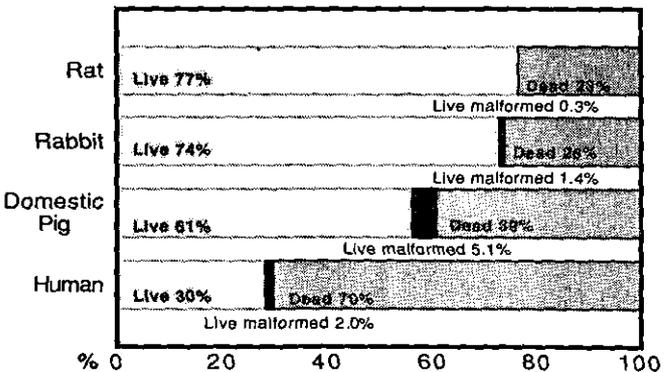


FIGURE 3. Outcome of pregnancy.

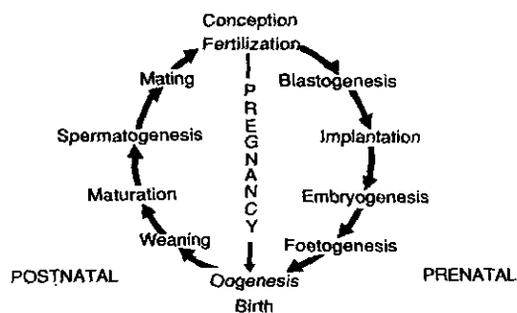


FIGURE 4. Cycle of life/reproduction.

prior to mating of mature adults through puberty of the offspring. Observations would need to be continued through conception in the second generation for detection of latent manifestations. In other words, testing requires a combination of exposure and observations through one complete reproductive or life cycle (Fig. 4). As we are dealing with a highly integrated process, only whole mammals, such as rats or mice, can be perceived as reliable surrogates for humans.

The existing test design that most closely fits these requirements is the two-generation study required for testing of agrochemicals and food additives. It would also appear to be the design most likely to be of use for testing environmental factors because it is intended for protection against involuntary exposure of large populations rather than voluntary use by the individual. Current versions of the different agencies unfortunately show irritating minor differences (Table 1), and they could be replaced (12) by a simpler, more practical design (Fig. 5). In this design the first generation could be exposed to the test conditions from shortly before mating: exposure and observations for various manifestations of reproductive toxicity are continued through the second generation.

Prenatal effects are detected by their impact on postnatal observations: this is also the principal of the Chernoff Kavlock assay (21), which has been proposed as a means of

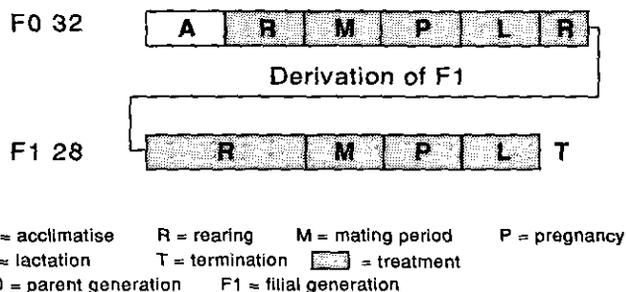


FIGURE 5. A simple two-generation study.

Table 2. Standard and nonstandard indices of reproduction.

Index	F <sub>0</sub> generation			
	1	2	3	4
Nonstandard indices, % <sup>a</sup>				
Genetic continuity	-33	-50	-54	-58
Weight weaned (total)	0	-39	-47	-65
Weight weaned (/dam) <sup>b</sup>	0	-18	-23	-44
Standard indices, % <sup>c</sup>				
Female mating	79	79	75	54
Gestation	100	88	92	85
Viability	97	97	98	88
Lactation	95	90	91	92

<sup>a</sup>Nonstandard indices more sensitive.

<sup>b</sup>Percentage difference from control values.

<sup>c</sup>Standard indices reset zero at each step.

detecting factors in the environment that might induce malformation. If need be, continued mating of one or both generations readily transforms this design into a fast breeding study (22). The fast breeding study has also been proposed as a means of detecting environmental factors harmful to reproduction. It has earlier, if forgotten, origins, as a test for the nutritional value of mouse diets.

The two-generation study is an apical test which, if interpreted correctly, makes it quite efficient at detecting whether an effect has occurred. Unfortunately, in the current climate of regulatory testing, the most popular methods of data analysis and interpretation leave some-

Table 1. Multigeneration studies: some comparisons.<sup>a</sup>

	Japan	FDA	EPA	OECD	Simple
Generations	2-3	2-3	2	2	2
Mates/generation	1-2	1-2	1	1-2	1-2
Premate weeks, F <sub>0</sub> M:F	8:2	10:2	8	10:2	2-4
Mating ratio preferred	1:1	1:1	1:1	1:1(1:2)	1:1
Standardization	Yes	Yes	Yes	Option	No
Random selection	Yes	Yes	Yes	Option	No
Routine histopathology	Yes	Option	Yes	Option	No
Different F <sub>1</sub> /F <sub>0</sub> size	No	Yes	No	No	Yes
Time to information of critical mass					
F <sub>0</sub>	≥20	22	≥20	22	14-16
F <sub>1</sub>	>42	>50	≥42	44	≤34

Abbreviations: FDA, U.S. Food and Drug Administration; EPA, U.S. Environmental Protection Agency; OECD, Organisation for Economic Cooperation and Development.

<sup>a</sup>Based on official written format, which is not always consistent with application in practice.

thing to be desired (Table 2). Further, the apical nature of the study that makes it efficient in detecting an effect also makes it difficult to determine the specificity and origin of any effect observed.

For more specific identification of effects, a segmental approach is preferred. Details of the three segment design studies are given in another paper presented at this workshop (25). As mentioned previously, the current three-segment design used for drugs is inefficient. Perhaps a way to better segmental designs is to re-derive them as subdivisions of a two-generation study.

As a step on the way, it can be seen that the current European Community (EC) segment 1 design (Fig. 6) is a two-generation study in which there is no direct exposure of the second generation. A second difference is that some females are killed before term so they may be examined for prenatal effects. These features weaken the power of the study more than they increase the specificity.

A first step to attaining a noticeable level of specificity is to subdivide either the two-generation study or an EC segment 1 study into two parts. One part emphasizes investigation for developmental toxicity, the other emphasizes investigation of adult fertility.

For investigation of developmental toxicity, exposure starts at implantation and continues through weaning. A second generation is reared to maturity to enable detection of latent effects. The design is almost identical to that of the Japanese experiment 3 but also incorporates exposures during organogenesis (Fig. 7).

For the complementary study for effects on adult fertility, exposure is initiated just before mating and terminated at or just after implantation. Females are killed and examined at about day 14 of pregnancy. This provides a fertility study equivalent to the current Japanese experiment 1 (Fig. 8).

In combination, these two parts cover all the exposures and observations required for testing pharmaceuticals, except for intensive examination for fetal abnormalities. As indicated earlier, I do not think this a serious omission. Objectively, the risk of inducing malformations is slight, and if it occurred there would be a high probability of detection by postnatal end points. Further, in contrast to the devastating impact of malformations on individuals, the impact on populations is negligible.

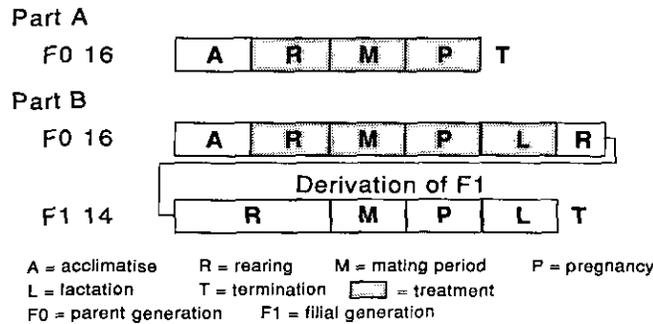


FIGURE 6. A simple fertility and general reproduction study similar to the current European Community segment 1 study.

However, for those who believe otherwise, one solution would be to extend the exposure period of the fertility study and conduct detailed fetal examinations for detection of abnormalities (Fig. 9). This design would provide a better match between observations and exposure periods than the current Japanese design and better group sizes than the corresponding part of the EC segment 1.

Another solution would be to leave the fertility study as it is and add the familiar, conventional segment 2 study (Fig. 10) to provide a true three-segment design. This

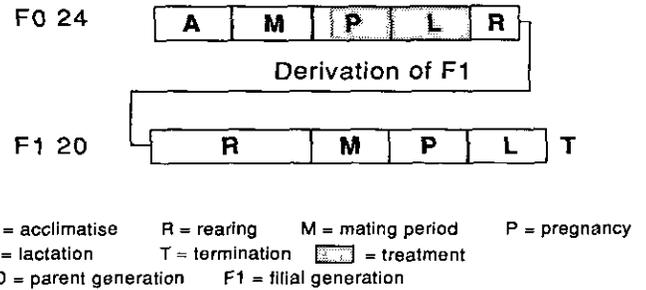


FIGURE 7. A pre- and postnatal (developmental toxicity) study equivalent to a Japanese experiment 3 study with an extended dosing period.

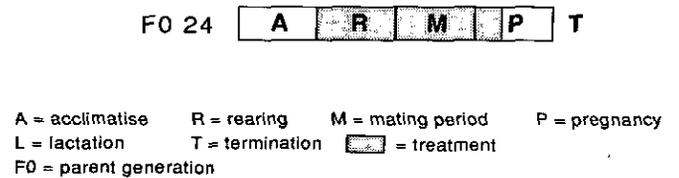


FIGURE 8. A fertility study similar to current Japanese experiment 1.

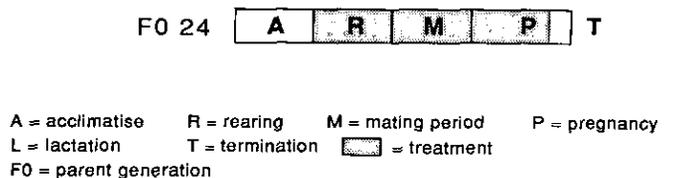


FIGURE 9. A fertility and embryotoxicity study.

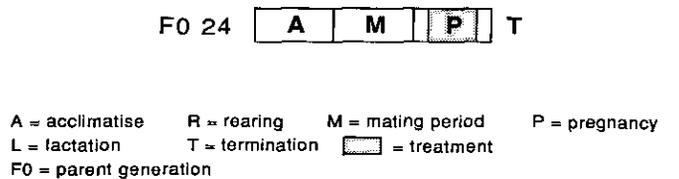


FIGURE 10. An embryotoxicity study identical to a conventional segment 2 study.

would be far less wasteful of time and animals than current 3 segment designs because unnecessary duplication of reared second generations and fetal examinations are avoided.

A major weakness in all past, present and proposed tests is that, irrespective of the duration of the pre-mating treatment period, mating is an insensitive means of detecting effects on males. A more sensitive method, especially for detecting effects on spermatogenesis, is the direct pathological examination of testes and accessory glands. Such examinations can be done in repeated-dose toxicity studies as well as in designated reproduction studies.

In overview, the various designs can be seen as a sequence starting with a single two-generation study, progressing through a two part subdivision, then a three-part subdivision, and beyond to case-by-case designs (Fig. 11). Through this sequence, the emphasis of testing changes from detection of any adverse effect through increasing specificity and characterization of effects.

## Application to Environmental Factors

Simplifying individual tests and linking them logically provides a role model for testing major new factors or strongly suspected existing factors. The focal point is a two generation study for detection of any effect. Should one be detected then the segmental subdivisions can be used for secondary evaluation and clarification. Results of testing suggest that secondary evaluation would not be frequent as only a small proportion of materials induce specific or selective effects on reproduction.

This still leaves the problem of dealing with unexplored factors in the environment. Given the enormous number of unexplored factors and the complexity of the environment, even the prospect of conducting a single, two generation study is daunting. There is a need for an even simpler methodology that at least can indicate the priority for testing. In this respect the test that provides the maximum information for the minimum effort is a one generation study involving exposure and observation of animals

from just prior to mating through weaning of the offspring; in other words, the first part of the two-generation study (Fig. 5).

Extending testing beyond a one-generation study greatly increases the time and effort, but produces very little extra information. Omission of the  $F_1$  generation and associated investigations for delayed manifestations may be a justifiable risk given the fact that such effects are rare. This is demonstrated by the low frequency with which effects are first detected in the second generation (23,24) of a two generation study (Table 3).

Regarding the preoccupation with teratogenicity, results from numerous Chernoff/Kavlock assays indicate that postnatal observations effectively detect teratogens indirectly. For example, a rat reproduction study (Table 4) provides unequivocal evidence that thalidomide would be harmful to reproduction when, in the same species, a standard "test for teratogenicity" does not.

Table 3. First detection of effects in multigeneration studies.<sup>a</sup>

Generation	Number of studies			Percent of studies		
	A	B	C	A	B	C
$F_0$	37	23	20	52	68	36
$F_1$	5	2	0	7	6	0
$F_2$	2	0	0	3	0	0
No effect	27	9	35	38	26	64

<sup>a</sup>A = Clegg (23); B = Palmer (12); C = Christian (22).

Table 4. Thalidomide effects in rat reproduction studies.

Mating	Litters born, %		Litter size born	
	0 mg/kg	200 mg/kg	0 mg/kg	200 mg/kg
Study 1				
1	79	32	8.3	3.5
2	74	17	8.9	4.7
3	83	6	8.8	3.0
Study 2				
1	50	5	9.2	4.0
2	62	0	9.5	—
3	75	0	9.8	—

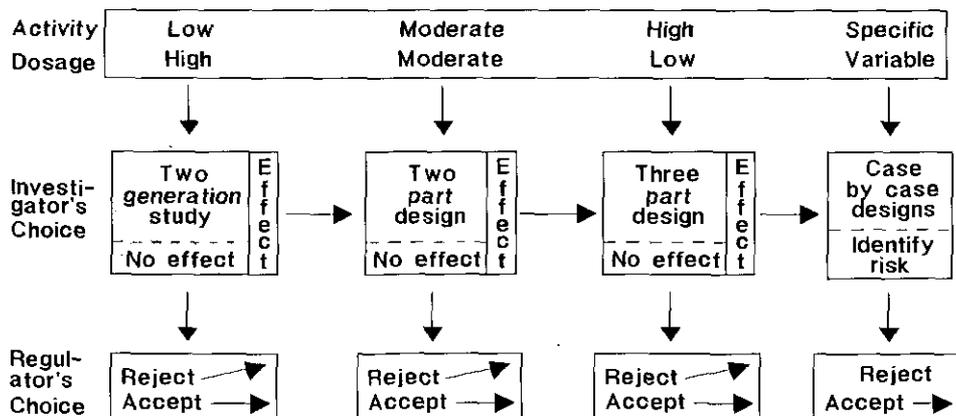


FIGURE 11. Reproductive toxicity: selecting studies.

Unfortunately, the most convincing testimony to the value of such a one-generation study is not openly available but hidden in company archives. Tests with the same or similar format are used by many reproductive toxicologists as preliminary studies to the studies required for testing chemicals and pharmaceuticals.

At levels of testing simpler than a short one-generation study, the situation is reversed because the saving in time and effort is negligible, whereas the diminution in scope and relevance to human reproduction increases enormously. This is true not only for *in vitro* tests but also for very similar Chernoff/Kavlock assay and for the familiar segment 2 (embryotoxicity) study.

## Summary

In summation, we need to use all of the methods mentioned to identify factors in the environment harmful to human reproduction. The expensive, time-consuming wildlife surveys and the reputedly cheap short-term tests for different reasons would appear to be limited to providing a source of hypothesis generation. A simple one-generation study would seem to provide the best means of selecting priorities for further testing. It might provide the best means for initial exploration of suspect multifactorial environments to which relatively large populations may be exposed.

Stripped of the layers of administrative extravagance and simplified, the testing methods now used for chemicals could provide a means of investigating new factors or strongly suspected existing factors. They provide the first point at which we might entertain the idea of presuming the absence of effect. Most certainly we need to put effort into human surveys as these provide the only certain way of identifying whether or not environmental factors will affect human reproduction. Whether there is a need to develop new methodology is doubtful: I suspect that we have methodology but lack the wisdom to use it wisely. We need better understanding of the environment and need to develop an appropriate philosophy in order to use existing methods within their limitations and in a balanced and objective manner.

## REFERENCES

- Holmes, L. B. Teratogen update: Bendectin. *Teratology* 27: 277-281 (1983).
- Brent, R. L. The Bendectin saga: another American tragedy. *Teratology* 27: 283-286 (1983).
- Brent, R. L. Teratogen update: Bendectin (editorial). *Teratology* 31: 429-430 (1985).
- Brent, R. L. Bendectin and interventricular septal defects (editorial). *Teratology* 32: 317-318 (1985).
- WHO. DDT and Its Derivatives - Environmental Aspects. Environmental Health Criteria 38, World Health Organization, Geneva, 1989.
- Newton, I. Population Ecology in Raptors. A. D. Poyser, Berkhamsted, 1979, pp. 229-252.
- Newton, I. Organochlorine insecticides and the sparrowhawk: toxic effects and population trends. In: *Ecotoxicology* (C. Walker and D. Anderson, Eds.), Institute of Biology, London, 1989, pp. 5-11.
- Palmer, A. K. Regulatory requirements for reproductive toxicology: theory and practice. In: *Developmental Toxicology* (C. A. Kimmel and J. Buelke-Sam, Eds.), Raven Press, New York, 1981, pp. 259-287.
- Alder, S., and Zbinden, G. National and International Drug Safety Guidelines. M.T.C. Verlag, Zollikon, 1988.
- Speid, L. H., and Lumley, C. E. Harmonisation of guidelines for toxicity testing of pharmaceuticals by 1992. *Regul. Toxicol. Pharmacol.* 12: 179-211 (1990).
- Manson, J. An overview and comparison of reproductive and developmental toxicity regulations. In: *Current Issues in Reproductive and Developmental Toxicity* (C. E. Lumley and S. R. Walker, Eds.), Key Publishing, Lancaster, 1991, pp. 21-34.
- Palmer, A.K. A simpler multigeneration study (poster presentation). In: *Proceedings of the 6th IUPAC Congress of Pesticide Chemistry* (R. Greenhalgh and T. R. Roberts, Eds.), Blackwell Scientific Publications, Oxford, 1986.
- Palmer, A. K. What is the ideal approach to reproductive toxicity testing? In: *Current Issues in Reproductive and Developmental Toxicity Testing* (C.E. Lumley and S.R. Walker, Eds.), Key Publishing, Lancaster, 1991, pp. 83-99.
- Lumley, C. E., and Walker, S. R. *Current Issues in Reproductive and Developmental Toxicity Testing*. Key Publishing, Lancaster, PA, 1991.
- Bass, R., Ulbrich, B., Palmer, A. K., Hildebrandt, A. G., Weissinger, J., Doi, O., Baeder, C., Fumero, S., Harada, Y., Lehman, H., Manson, J., Neubert, D., Omori, Y., Sullivan, F., Takayama, S., and Tanimura, T. Draft guideline on detection of toxicity to reproduction for medicinal products. *Adverse Drug React. Toxicol. Rev.* 9: 124-141 (1991).
- Lumley, C. E. Proposal for international guidelines for reproductive and developmental toxicity testing for pharmaceuticals. *Adverse Drug React. Toxicol. Rev.* 10: 143-153 (1991).
- EPA. Air Quality Criteria for Lead: Executive Summary and Conclusions, Vol. 1. U.S. Environmental Protection Agency, Research Triangle Park, 1986, pp. 154-155.
- Herbst, A. L., and Anderson, D. Clear cell adenocarcinoma of the vagina and cervix secondary to intrauterine exposure to diethylstilboestrol. *Semin. Surg. Oncol.* 6: 343-346 (1990).
- Ehrhardt, A. A., Meyer-Bahlburg, H. F., Rosen, L. R., Feldman, J. F., Veridiano, N. P., Elkin, E. J., and McEwen, B. S. The development of gender-related behaviour in females following prenatal exposure to diethylstilboestrol. *Horm. Behav.* 23: 526-541 (1989).
- Hardin, B. D. Evaluation of the Chernoff/Kavlock test for developmental toxicity. *Teratog. Carcinog. Mutagen.* 57: 1-127 (1987).
- Lamb, J. C., and Chapin, R. E. Fertility assessment by continuous breeding. *J. Am. Coll. Toxicol.* 4: 173-184 (1985).
- Christian, M. S. A critical review of multigeneration studies. *J. Am. Coll. Toxicol.* 5: 161-180 (1986).
- Clegg, D. J. Animal reproduction and carcinogenicity studies in relation to human safety evaluation. In: *Toxicology and Occupational Medicine* (Diechman, Ed.), Elsevier, Amsterdam, 1979, pp. 44-59.
- Kimmel, C. A. Approaches to evaluating reproductive hazards and risks. *Environ. Health Perspect.* 101(Suppl. 2): 137-143 (1993).