Experimental teratology

IV. EXPERIMENTAL TERATOLOGY

A. Basic principles

The basic principles of teratogenic drug testing are similar to those underlying the detection of toxic reactions in adults. Therefore drug teratogenicity can be considered an aspect of general toxicology. However the actions of drugs on the embryo are more complex than those on the adult organism. In teratogenesis one is dealing with two biological systems – the pregnant female, and the embryo – the specific reactions of which can be completely different. Consequently a drug that is non-toxic to the female may still be apt to kill the embryo or produce congenital malformations. The problem is even more complex since the influence of the placental transfer of drugs and their metabolic fate in the embryo are only partially understood. To compete with the various mechanisms which control prenatal development, very complex conditions have to be fulfilled.

The action of a teratogenic agent on the conceptus depends mainly on three conditions: (1) the developmental stage of the embryo, (2) the genetic susceptibility of the embryo, and (3) the physiological or pathological status of the mother.

1. The developmental stages

The period in which injurious agents mainly affect the development of the human embryo is very short. It is largely completed by the 8th week of pregnancy, just about the time when a women knows that she is pregnant (Fig. 4).

During the pre-implantation period, when the blastocyst lies free within the uterus and depends for its nutrition on the uterine secretions, exogenous agents can kill the embryo, but there is no evidence that they can produce congenital malformations. This is the period of maximal embryolethality. However, minor injuries can be overcome without manifest harmful consequences on the growing embryo because, during the segmentation stage, many blastomeres retain their totipotency, being able to replace damaged cells by newly formed cells. Once implantation has occurred in the human, 7–8 days after fertilization, the embryo undergoes very rapid and important transformations.

The sequence of embryonic events shows that each organ and each system undergoes a critical stage of differentiation at a precise moment of the prenatal development. It is during this critical period that the vulnerability of the developing embryo is greatest and that specific gross malformations can be produced. *This is the susceptible period*, which lasts until the 56th day in the human embryo. *The foetal period* begins at the end of the 8th week, when little further differentiation of organs remains to be completed. The most important events at this stage are the closure of the palate, the reduction of the umbilical hernia at the end of the 9th week, the differentiation of the external genitalia, as well as the histogenesis of the central nervous system.

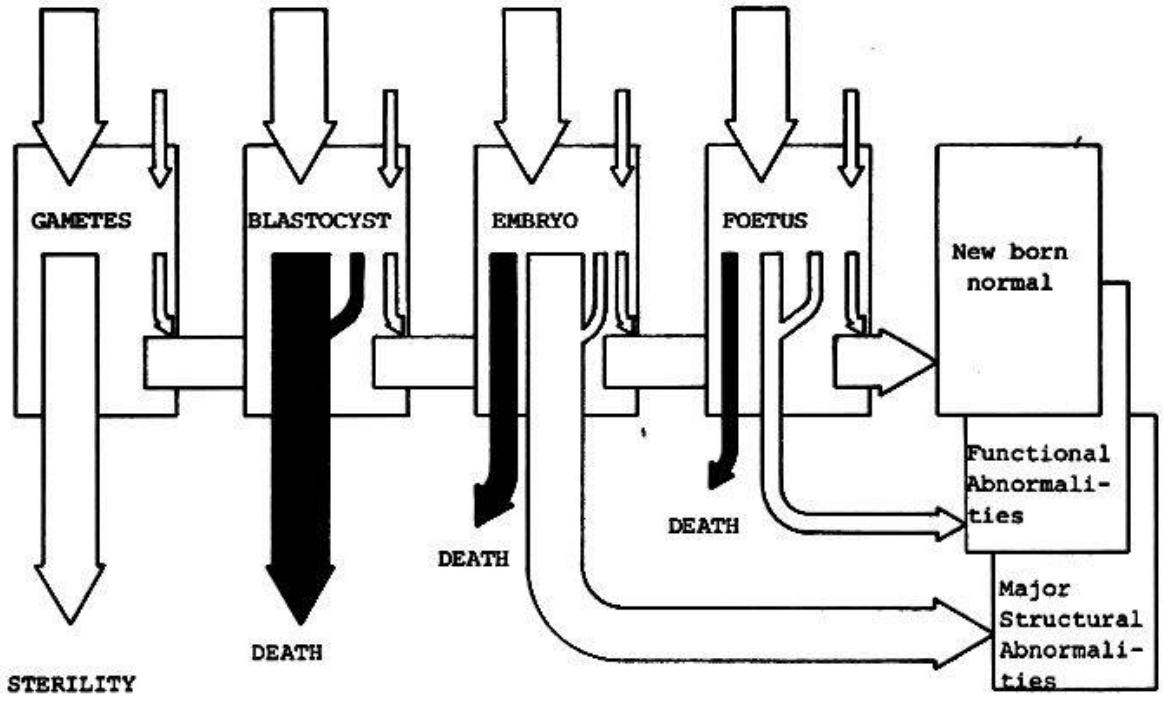

Fig. 4. Schematic representation of the influence of teratogenic factors on gametogenesis and various stages of prenatal development. Strong teratogenic agents: wide arrows; weak teratogenic agents: narrow arrows.

This latter process lasts for the entire period of intra-uterine development and is not complete until several months after birth. Consequently, during the foetal period, teratogenic agents do not produce major morphological malformations but can impair the differentiation of external genitalia, leading, in severe cases, to pseudohermaphroditism. Interference with the histogenesis of the central nervous system can lead to various degrees of encephalopathy.

2. Genetic susceptibility and species differences

The reaction of the embryo to exogenous agents depends upon its genetic constitution. The reaction of the embryo to a specific compound varies not only between different animal species, but also within a given species, between each strain, and even between individuals of the same strain (Figs. 5 and 6 may serve to illustrate the differences in the critical stages of embryogenesis between two species of laboratory animals).

Cortisone, a potent teratogenic agent in the rabbit and in the mouse, does not produce malformations in the rat. Thalidomide, which produces obvious malformations in the rabbit, is apparently innocuous in the rat. A purine analogue, azathioprine, which is highly teratogenic in the rabbit and the mouse, does not produce anomalies in the rat. Teratogenic susceptibility differences in strains have been demonstratcd. For example, the same regimen of treatment with cortisone causes 18% malformations in C57 BL mice while in the A/Jax strain 100% malformations are produced.

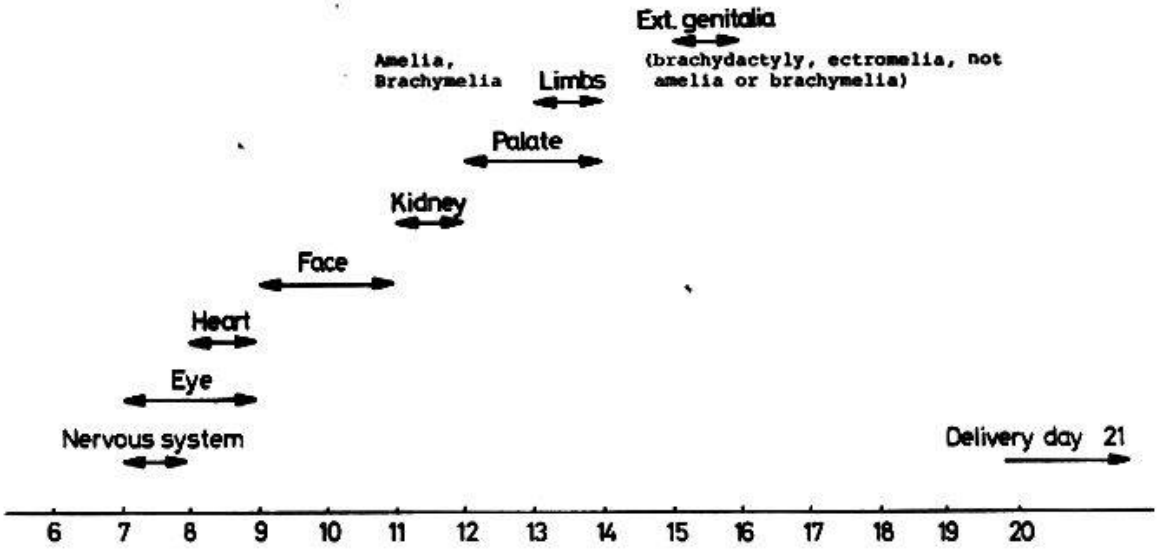

Fig. 5. Critical periods in embryogenesis of the rat.

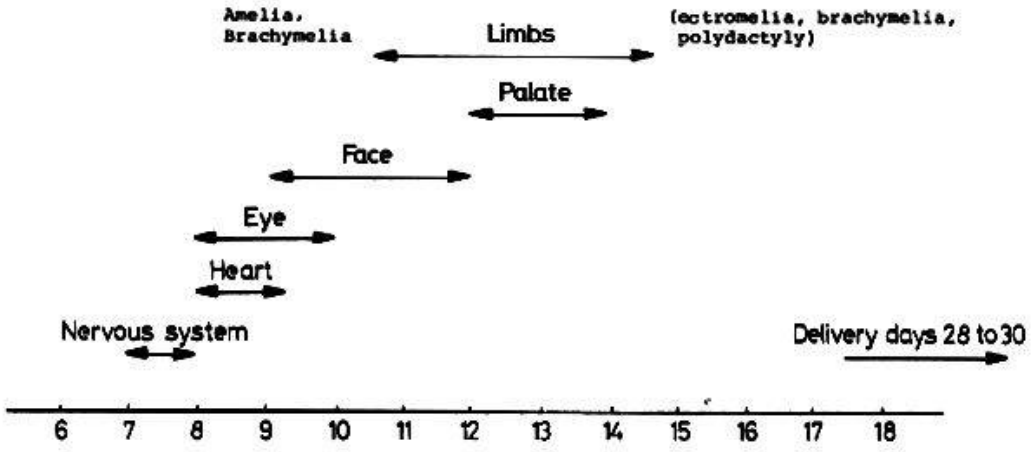

Fig. 6. Critical periods in embryogenesis of the rabbit.

3. Physiological or pathological status of the mother

Besides the developmental stage and the genetic constitution of the embryo, the action of a drug is dependent on the physiological and pathological condition of the mother. The most important physiological factors are: age, diet, hormonal balances, and uterine environment. Among pathological conditions, metabolic diseases, diabetes and obesity tend to increase the action of potential teratogenic agents.

B. Teratogenicity testing

1. The experimental procedures

In the determination of the toxicological potential of a drug, the choice of the conditions of treatment, that is to say, the route of administration, the doses, the number of animals per dose, the duration of treatment, etc., is of major importance.

The problem is particularly difficult in teratology because of the complex nature of the biological system that is involved (see chapter III).

In addition, it must be noted that the treatment may also, under certain conditions, (excessive doses, local irritating effect at the site of injection, etc.), induce in the mother pathological changes which by themselves are incompatible with the survival and the normal development of the embryo: such conditions should be avoided, since abnormalities observed in the foetus would be devoid of any specificity.

The toxicity of the compound for the embryo may manifest itself either by the death of the litter or by the appearance of malformations. A compound may cause embryolethality without being necessarily teratogenic; but with some drugs, these two toxic effects are associated and, in such a case, the early death of the embryo, followed by resorption, may hide a teratogenic effect: the teratogenic activity may then be detected only in a rather narrow range of doses, at which embryolethality remains still moderate.

The sensitivity of the embryo to the toxic effects of a compound depends on the developmental stage reached by the embryo when it is submitted to the action of the latter. One may consider that the stage of embryogenesis properly speaking (or organogenesis stage), which begins in most species at the time of blastocyst implantation and takes place between the blastogenesis stage and the foetal development phase (or foetogenesis phase), corresponds to the period of sensitivity to teratogenic effects. If a toxic injury manifests itself during the formation of the blastocyst, there are, it appears, with very rare exceptions, only two possibilities: the blastocyst dies, or it overcomes the injury, the consequences of which for the subsequent development of the embryo are then very limited or non-existent.

Furthermore some authors suggest that if virgin females of the rodent species are used, and if the treatment is begun after the implantation, which has the advantage of eliminating a possible effect of the compound on the implantation process, it is easier to check, *a posteriori*, the pregnant state by detecting the implantation sites when sacrificing the mothers before delivery.

Although in teratological studies, as in any type of toxicological study, the investigator must be allowed to adjust his protocol to the particular case of the compound under study, and not be curtailed by too rigid rules, it is possible to propose some general recommendations for the conditions of treatment. First of all, the state of health and standards of care of the animals must be carefully controlled. Animals that are to be used for teratological studies of drugs must be healthy and should be housed under the best possible environmental conditions. The best practices of animal care must be maintained; specified pathogen-free animals are not necessarily required under these conditions. It is advisable that strains with known genetically unstable constitutions be avoided. The animal quarters should provide constant temperature, adequate light and protection from noise or other interference. When temperature and light cannot be controlled, the seasonal variation of the reproductive activity of the animals must be considered.

Caging should conform to the best standards available. The use of pesticides must be avoided. It must be remembered, particularly with drugs that act upon the central nervous system, that housing groups of animals in a single cage may increase the noxious effects of certain drugs (group toxicity). All animals should receive an adequate diet; since the administration of a drug may decrease the food consumption of the treated animals, it may be necessary to adjust the food intake of control animals correspondingly.

a) Number of species

For routine teratogenic drug screening animals with a high fertility rate and a short duration of pregnancy should be used. At least two species, one rodent species, the rat or the mouse, and one lagomorph, the rabbit, will fulfil the general requirements.

b) Choice of route and method of administration

Generally the drug should be administered by the route intended to be used in man.

In the case of the oral route, one must determine (preferentially by blood-level determinations) whether the compound is absorbed by the species under study at the level of the gastro-intestinal tract. If the compound is insoluble in the chosen vehicle, its particle size must be sufficiently small and well defined, in view of the role this parameter plays in the digestive absorption.

The compound may be administered in the food on the strict condition that its stability in the food has been ascertained. The amount of the compound ingested will be estimated on the basis of the daily food intake. In the view of some authors this method of administration might have the advantage of covering all the developmental stages by more or less consistent exposure of the target organs to the compound and/or to its metabolites. However, the administration of the compound by stomach tube (either as a solution or a suspension in the proper vehicle), particularly in the mouse and the rat, or in capsules (in larger animals), seems to be more frequently used. It has the advantage of allowing the amount of the compound ingested to be accurately determined and usually ensures high levels of absorption.

The use of the intravenous, intramuscular, or subcutaneous routes is generally without problem but the interpretation of the results may be difficult, when the local tolerance at the site of injection is poor. It appears preferable to avoid the intraperitoneal route, because of the risk of accidental injection at the level of the uterus and of the effects of a possibly poor local tolerance.

Although it may be conceivable to study the teratogenic effect of a compound administered by the rectal or vaginal routes or by inhalation, the use of these routes is fraught with difficulties, especially in the smaller animal species (problem of checking that the suppository remains in the rectum; risk of abortion by vaginal stimulation; hypoxia and stress caused by some inhalation procedures).

In the case of a compound to be used in man by one of these routes, it appears possible to take into consideration the results of an experiment performed by the oral or, better, by the intravenous, intramuscular or subcutaneous routes. A prior comparative study of the metabolism of the compound in the animal species used, administered by the route intended for therapeutic use and by the route chosen for the experimental study, will generally justify the decision taken.

For compounds to be used for dermal, nasal, auricular, or ocular treatments, a study by any route that is demonstrated to result in systemic exposure may be adequate.

Finally, during the course of treatment females should not be subjected to any kind of rough handling, which by itself could cause disturbances of the pregnancy.

c) Period and duration of treatment

As a general rule, the treatment must begin immediately after implantation and continue during the whole period of organogenesis. In experiments using rodents, for instance, the following periods appear to be adequate:

- in the mouse and the rat, from the 6th day to the 15th day inclusive, day 0 being marked in the mouse by the presence of a vaginal plug, and in the rat by the presence of spermatozoa in vaginal smears;
- in the rabbit, from the 6th day to the 18th day inclusive, 0 day being that of mating.

Nevertheless, in a number of cases, e.g. severe impairment of the health of the mother animal caused by excessive pharmacological activity, tolerance developing to the compound tested owing to the induction of metabolizing enzyme, etc., it may be useful or even necessary to limit the treatment of the pregnant females to a few days or even a single day, provided the whole period of organogenesis is covered by using an appropriate number of animal groups.

Furthermore, the pharmacokinetic characteristics of the compound in the animal species that is tested should be taken into account. In the case of a compound with a short half-life, for instance, administration of the compound twice a day during the whole period of organogenesis may be envisaged. Alternatively, in the case of a compound with a long half-life period, slow release or depot preparations in particular, intervals of several days between each consecutive administration may be appropriate.

d) Choice of doses

The selection of adequate doses is most important.

In the case of drugs the factors to be taken into account are the following:

- the nature of the pharmacological activity exerted in the particular species and the dose at which this activity manifests itself
- the therapeutic dose of the compound in man

- pharmacokinetics and metabolism in the animal species in comparison to man
- the toxicity of the drug to the mother animal

e) High dose effects

Since the aim of toxicity studies in general is to detect untoward effects and to demonstrate how these effects are related to the dose, high doses up to the maximum tolerated dose are commonly administered.

This procedure has its limitations in the case of certain drugs such as hypnotics, local anaesthetics, hypotensives and diuretics, for instance, which, owing to excessive pharmacological activity, may impair the general condition of the animal including metabolism, circulation, nutrition and temperature regulation, secondarily leading to organ damage. The developing foetus is particularly sensitive to this kind of disturbance and early foetal death, resorptions, malformations, or reduction of growth may occur.

Very high doses which might impair the general health of the pregnant animal can produce malnutrition of the embryo and consequently the appearance of an unspecific teratogenic effect: such an effect, although presenting difficulties in the interpretation of the results of the teratological study, would have no significance with respect to the normal conditions of use of the drug in clinical practice. Furthermore, by administering excessively high doses, without any relation to the pharmacological doses, the metabolism of the compound may be modified quantitatively as well as qualitatively.

With these considerations in mind it appears desirable that the highest dose studied should exert in the pregnant female some systemic effect, for instance, a slight impairment of the weight increase in comparison with control animals. It is important that the highest dose does not cause excessive embryolethality and the foetal loss should not be higher than 50%.

Teratogenic effects produced by extremely high doses, in comparison with the human intake, demand special consideration. In general, they are considered relevant only if these doses did not impair the well-being of the mothers and if they were not dose-dependently accompanied by other disturbances of the pregnancy.

f) Preliminary investigations and final experiments

The choice of the maximum tolerated dose makes it desirable to perform a preliminary experiment on a few pregnant females, with 2 to 3 doses and a limited number of animals per dose.

Such a study is considered necessary because the pregnant female is usually more susceptible to the general toxic activity of a compound than the normal female.

In the final experiment, one should add to the highest dose lower dose levels. In order to establish a dose-response relationship generally two doses are chosen on a logarithmic scale: the lowest dose, which must not cause any embryolethality, will approach the therapeutic dose and is usually chosen

to be twice or three times higher than the dose (per kg of body weight) intended for use in man.

A series of control animals will be treated with the vehicle, under the same conditions as the animals receiving the compound under study. The number of animals per series must be sufficient to enable a statistical evaluation of the results to be made (cf. Chapter IV, B 4).

g) Choice of the animal species

Rodents. A large variety of malformations can be experimentally induced in rodents. In the mouse the most frequent gross malformations involve the palate, the central nervous system and the tail. In the rat, malformations of the eye and skeletal deformities may predominate.

In lagomorphs, eye malformations are quite exceptional whereas coelosomia and central nervous system and skeletal anomalies have been among those most frequently observed. However it must be emphasized that there are no general rules. Usually there is a relationship between the morphological type of the anomaly and the time of drug treatment.

Rat. Often considered as a reference animal, the rat has several advantages over other species: short duration of pregnancy, high fertility rate, large litters, and relative resistance to the toxic effects of many compounds. Rats also have a fairly good developmental stability. In the Wistar line the spontaneous rate of gross malformations is approximately one per 1,000 foetuses. The main limitation of the rat is its poor teratogenic susceptibility to some drugs, like cortisone, thalidomide, and azathioprine, for example. It is therefore not recommended to draw definite conclusions on the basis of results obtained only in this species.

Mouse. This species shares the advantages of the rat and in addition, mice are particularly susceptible to some teratogens. In Swiss albino mice the spontaneous malformation rate is approximately 0.5%. The resorption rates are also higher than in the rat. Therefore it may be necessary to have a large control group.

Other rodents have been used only rarely for teratogenic investigations.

The few teratological investigations performed *in hamsters* have given satisfactory results comparable with those obtained in rats and mice. However, from the available data it does not appear that this species offers any advantage for teratogenic drug testing.

In *guinea-pigs* pregnancy lasts three times as long as in rats and litters contain only 2–4 foetuses. It is desirable to obtain more information on this species.

Lagomorpha. The discovery of the susceptibility of rabbits to thalidomide led many biologists to consider them as one of the most favourable animals for teratogenic studies. The use of the rabbit has also been advocated as a non-rodent species.

The pig. The possibility of using the common domestic pig (Sus scrofa) for teratogenic screening has recently been explored. The pig is far removed in

evolutionary descent from rodents and could therefore fulfil official requirements of a non-rodent representative. The chief advantages of the pig as compared to other non-rodents are that it is easily available and highly prolific (average litter 10–11). Oestrus occurs at 21-day intervals and mating can take place at 7–8 months of age. The gestation period is about 115 days, which is of some advantage when compared with the rhesus monkey, which has a gestation time of 164 days. The embryology and the genetics of the pig are fairly well known. The incidence of spontaneous malformation is relatively low. The susceptibility of pig embryos to teratogenic drugs seems to be fairly high; vitamin A induces teratogenic effects and thalidomide causes a high incidence of visceral, facial and urogenital malformations.

In experimental practice, however, the use of the pig is much limited. The expenditure of care, including the large amount of space and the comparatively enormous amounts of compound required to perform toxicological and teratological studies, constitute real handicaps. It was hoped that the problem of size would be overcome by the use of a miniature breed of pigs, but minipigs are not readily available and their fertility and prolificacy are low. More experience is needed before a statement can be made as to the usefulness of the pig in teratological investigations.

The dog. Although polytocous like rodents, dogs have a recurrent oestrus at intervals of about 6–8 months and even longer when maintained under laboratory conditions. Induction of ovulation by hormonal stimulation is not advisable owing to the possibility of drug interaction. The limited period of heat and the long gestational period require a considerable amount of time to perform teratogenic studies. Further, the interpretation of the experimental results requires data on the normal incidence and type of malformation for the particular breed used.

As far as the teratogenic susceptibility of dogs is concerned it is difficult to make a definitive statement because only limited data are available. The critical period of organogenesis seems to take place between days 12–27.

The cat. Very few data are available for this species, which is rather difficult to maintain and breed under laboratory conditions. The most favourable period for drug-induced teratogenicity seems to be between days 12–27 of gestation. Particularly the difficulties encountered in breeding are against the use of cats for routine teratogenicity testing.

Primates. The phylogenetic proximity of monkeys and human beings and the great similarity of reproductive physiology in women and certain monkeys, like the rhesus, led to the suggestion that this species might be more reliable for teratogenic drug testing than other laboratory animals.

Wilson (1971) proposed a teratological test using 25–30 pregnant monkeys. He recommended the following procedure to find doses and treatment times causing embryotoxicity. A preliminary screen is conducted with 15 animals, one treated at each of three dose levels at five gestational stages, followed by hysterotomy on day 100. Larger doses are used if no embryotoxicity (growth retardation, intra-uterine death, or malformation) is observed

initially. Additional animals are then treated with doses and at times associated with embryotoxicity, to define these variables better, and some pregnancies are allowed to go to term if needed.

For a general appraisal of the possible usefulness of monkeys in teratogenic testing, the main results obtained in recent years need to be confirmed. Since several species of macaque monkeys and baboons react to thalidomide in the same way as man, it has been assumed that simian primates would be better and safer test animals for the evaluation of drugs than the conventional laboratory animals.

This assumption has not been substantiated by screening various drugs in rhesus monkeys. From his personal experience WILSON (1973) concluded "it must be emphasized that the close parallel between man and monkey as regards the teratogenicity of thalidomide has not been demonstrated for other drugs except those with androgenic properties".

Several drugs which are teratogenic in rodents, like meclizine, 6-aminonicotinamide, acetazolamide, methotrexate, etc. did not induce malformations in monkeys.

An interesting case is that of methotrexate, a folate antagonist, which is highly teratogenic in rats at 0.3 mg/kg (WILSON, 1970). In rhesus monkeys no teratogenic effects were observed with 3 mg/kg and even higher doses.

From his comparative observations WILSON (1971) concluded that rats are teratologically more susceptible to acetazolamide, retinoic acid, methotrexate, 5-fluorouacil, hydroxyurea and vincristine. Conversely, monkeys were particularly susceptible to thalidomide. The reason for these species differences is still not clear. It was suggested that it might be related to differences in the metabolism of test substances. However for a compound like methotrexate the available data on distribution and excretion do not reveal differences between rats and monkeys that could explain the differences in embryotoxicity. As pointed out by WILSON (1971) these data raise the question which of these two species is more appropriate for the teratological testing of drugs in general.

2. Particular problems raised by certain drugs

The experimental difficulties encountered in routine teratogenic drug screening include some problems inherent in the action of certain types of drug, for example:

Compounds which destroy the intestinal flora, as do certain antibiotics in the rabbit; drugs with a prolonged, intensive pharmacological activity; compounds causing enzymatic induction; compounds which rapidly harm the health of the mother and can be administered only for a short period of time. To overcome these difficulties, the treatment may be limited (cf. Chapter IV, 1c).

Psycho-active drugs

Substances which are capable of modifying the pituitary-hypothalamic system and of causing release of lactogenic hormone (butyrophenones, reser-

pine, phenothiazines, etc.) are likely to result in delayed implantation. Consequently the embryos exhibit apparent growth retardation which in fact reflects only the retarded implantation. Such a reaction is not likely to occur in humans.

Similar problems are also encountered with certain steroids or pituitary hormones when tested in conventional laboratory animals.

Hormonal agents

Corticosteroids: The catabolic activity of these preparations induces weight loss in the mother animal.

The suppression of the immune response, a well-known effect of corticosteroids, lowers the resistance to infections.

Glycogen accumulation and fatty infiltration in the liver and, in some species, liver necroses are produced.

The particular properties of corticosteroids influence the state of health of the mother animal and consequently may harm the foetus. With ACTH the same problems arise owing to stimulation of the adrenal gland.

It might be useful to perform the experiments in specified pathogen-free animals.

Many of the anti-inflammatory steroids produce cleft palate in the foetuses of animals dosed during pregnancy. The mouse is the most sensitive species and practically all anti-inflammatory steroids will induce cleft palate. The rat and the rabbit are far less sensitive but respond to the more potent glucocorticoids such as dexamethasone and triamcinolone. Particularly in mice this effect is not considered to be specific. Cleft palate can be produced in this species by treating the dam with ACTH, which suggests that mice are sufficiently susceptible to be affected by still physiological quantities of corticoids released by the adrenals. By contrast, adrenal stimulation by extreme stress fails to induce cleft palate in the rat. Reduction of amniotic fluid caused by amniocentesis is sufficient to induce cleft palate in mice. Cortisone is known to produce oligohydramnios and it is suggested that the higher sensitivity of mice to cleft palate production is related to their particularly rigid craniospinal articulation resulting in less craniospinal flexibility *in utero*.

In spite of many thousands of women having received cortisone during pregnancy there is no convincing evidence that this treatment results in cleft palate or other congenital malformations in man. It appears certain that the risk of teratogenic damage must be extremely small.

Sex hormones

In rodents and in rabbits oestrogens produce a weight loss and, depending on the dosage, partial or complete resorption of the embryo. In these species oestrogens control the uterine growth and the elaboration of uterine fluids.

Progestational compounds have no adverse effects on pregnancy even at high doses, provided that administration is restricted to the period of morpho-

genesis. In later stages the gestation can be prolonged and foetal growth and parturition are delayed. This would also be true for gonadotrophins.

Simultaneous administration of progesterone and oestrogens, as in oral contraceptives, moderates the oestrogen effect. Much higher doses are necessary to cause abortions.

Anti-oestrogens will cause abortion in much the same way as oestrogens. Anti-oestrogenic compounds like clomiphene and methallibure have the same effects in rodents and are also teratogenic, probably through a pituitary-blocking action. However these effects are not likely to occur in humans because there are important physiological differences.

The information gained in rodents, which have a particular hormonal control of gestation, raises the problem of the selection of a suitable animal model for the study of the reproduction processes as they occur in humans. For instance, it has been suggested that experiments be conducted in dogs and monkeys to test the chronic toxicity of fertility-regulating agents. This recommendation can be misleading as far as the use of dogs is concerned, because this species has a particular susceptibility to some progestational compounds and presents an exceptionally high incidence of spontaneous mammary tumours.

While in rodents and in rabbits the maintenance of pregnancy requires functioning ovaries, in women, in primates, and in guinea-pigs, the placenta takes over very early the hormonal control of pregnancy. Therefore, as in women, the maintenance of gestation is not dependent on active corpora lutea but on the placental activity. In conventional laboratory animals oophorectomy causes abortion during the major period of the pregnancy. In primates, after the first six weeks, pregnancy is maintained by the large amount of chorionic gonadotrophins, progesterone and oestrogens produced by the placental tissue. By contrast, in the rabbit the production of progesterone and oestrogen in the corpus luteum is governed by the pituitary; the placenta does not produce pregnancy-maintaining hormones.

3. Criteria of congenital malformation, visceral and skeletal examination[1]

Congenital malformations can be defined as "structural" defects present at birth. They can be external, internal, grossly or only microscopically detectable. Although substances may be well tolerated by the dam they can be embryotoxic and/or teratogenic.

Embryotoxicity may be subdivided into:

- Embryolethality (post-implantation death)
- Retardation of growth and development

Retardation can consist in inhibition of the physiological growth of the foetuses or in delayed development. Inhibition of growth exists when the

[1] Although the present emphasis is placed on morphological criteria, the possibility of "metabolic malformations", i.e. enzymatic defects without structural abnormalities of the tissues should also be considered.

foetuses are normally developed but are significantly under weight. Delayed development is indicated primarily by retardation of physiological ossification of the skeleton.

An impairment of the placenta can also indirectly impair the development of the foetuses. Thus it may be advisable to examine and weigh the placenta. However, experience has shown that examination of the placenta does not offer any practically relevant information.

The following suggestions regarding the examination of foetuses for skeletal and visceral anomalies apply principally to screening studies. Therefore, priority must be given to the adequate examination of a large number of foetuses for visceral and skeletal defects.

The method(s) of examination will depend on several factors including the facilities available, the preferences of the investigator, and the size of the foetus to be examined.

External examination

Thorough external examination of the foetus in utero is a prerequisite for the success of subsequent skeletal and visceral examinations.

Large species. For larger species such as dogs, monkeys, pigs, cats, etc. the individual foetus can be examined for both skeletal and visceral defects, a procedure which is perhaps essential in view of the tendency to examine small numbers in such species. For reasons of logistics and economy, radiography of the foetus immediately after external examination is probably the most convenient method.

For skeletal examination, X-rays should be taken in at least two planes. It must be remembered that, even with quite sophisticated apparatus and the higher degree of skeletal development of large species, radiography does not provide the precision and three-dimensional flexibility afforded by alizarin staining and some expertise is required to distinguish anomalies from artefact.

Following radiography all foetuses can be examined for visceral defects and probably the simplest and yet the most efficient method is to follow a standard autopsy procedure. It is important to dissect the urinary tract and to examine the heart, particularly the interventricular septum, and the main vessels. The intact brain should be carefully removed *in toto* and sliced in order to detect internal defects.

The standard autopsy technique is most efficient when performed on freshly killed foetuses (i.e. before fixation) as the coloration of the tissues allows ready identification of organs, the flexibility of the tissues allows them to be displaced and viewed from various angles, and the patency of blood vessels can be checked either by the simple application of pressure or by injecting a coloured fluid. On fresh specimens ocular defects can be detected more readily than after fixation.

Intermediate size species. For intermediate sized species, such as the rabbit, or larger species delivered at an early stage of pregnancy, skeletal and vis-

ceral examination can be performed along the lines indicated for large species. However, instead of radiographic examination the alizarin-staining technique would appear to be the method of choice of skeletal examination.

Visceral abnormalities may be rather easily detected by examining the body cavities, i.e., the thoracic and abdominal organs *in situ* (including pelvis, uro-genital tract).

There are different methods of examination of the various topographical and organ systems; all have their advantages and disadvantages. There is no method that allows all of the foetuses to be examined for both visceral and skeletal anomalies.

Small species. Although it is possible to examine small foetuses for both skeletal and visceral anomalies in a manner similar to that used for the larger species, the process is difficult and time-consuming; it requires a combination of dissection under a stereomicroscope and the use of alizarin staining or sophisticated radiographic apparatus.

Most laboratories, therefore, circumvent the difficulties of examining smaller foetuses by increasing the group size (relative to larger species) and apportioning foetuses so that some are examined principally for visceral defects whilst others are examined principally for skeletal defects.

Of the two methods of skeletal visualization, alizarin staining is preferable. The method usually followed is that of DAWSON (1926). In evaluating the results, due account must be given to the fact that ossification of various skeletal parts is particularly active at the perinatal period (FRITZ and HESS, 1970).

Foetuses allocated for skeletal examination can also be examined to a limited extent for major visceral abnormalities. One method is to make a transverse cut across the abdomen and a longitudinal cut (on one side of the sternum) from the neck to the inguinal region; the epidermis can then be turned back as flaps and the visceral contents examined *in situ* before being removed. Alternatively, one may make one transverse cut across the abdomen and, with a pair of fine forceps inserted up through the thorax, remove the viscera intact.

Although this technique will reveal major visceral malformations and occasional minor anomalies special techniques have been developed for the analysis of visceral defects.

The most commonly used technique (WILSON, 1965) consists in cutting a series of transverse sections (0.5 to 1 mm thick), principally through the head, thorax and abdomen, which are then examined under the microscope. Fixation of the foetuses must be rapid and as well as decalcifying the foetus must leave it sufficiently firm to be sliced thinly without causing disintegration of sections. If fixatives containing picric acid (Bouin's fluid) or formalin are used it is desirable to transfer the foetuses to another preservative (e.g. ethanol) before handling and slicing them.

An alternative microdissection was proposed by BARROW and TAYLOR (1969). This technique, although adequate, may not be quite as efficient as

the Wilson technique for detecting minor heart anomalies and, moreover, tends to preclude the opportunity of performing subsequent histological examination. However, it is a useful technique for supplementary examination of foetuses principally allocated for skeletal examination.

Of the various modifications of these two techniques one that appears to be of value consists in removing the heads from all foetuses to examine them by the Wilson technique, the caracasses being apportioned to skeletal staining and/or visceral examination.

Apportioning of foetuses and recording of results

Some government agencies suggest that 50% of small foetuses be examined for skeletal defects and the remaining 50% be examined for visceral anomalies. Other authorities suggest two thirds to be subjected to skeletal staining and one third to be examined for visceral anomalies.

Preferably more foetuses should be allocated for skeletal staining, for the following reasons:

- They can be examined for major visceral abnormalities before processing.
- Differentiation or development of the skeleton (or its precursors) continues throughout gestation and general effects elicited at any time during pregnancy may be manifest as a skeletal change.

Records should be kept in such a way that abnormalities can be related to the individual foetus, the individual litter and the technique of examination. Results should be reported as the absolute proportions of foetuses affected since percentage values can be misleading, particularly when they take no account of litter effects or the occurrence of several anomalies in one foetus.

4. Number in experimentation

Numbers of animals

When rodents are used the number of animals can be large enough to satisfy statistical requirements. Although this seems reassuring it must be borne in mind that to fulfil this requirement one has to define:

- the parameter(s) to be examined, its natural incidence and frequency distribution (which determines the most appropriate method of statistical analysis).
- the magnitude of the difference from "normal" which the analysis is required to detect.
- whether the difference will be examined by a one- or a two-tailed criterion.
- the level of significance (probability) that will be acceptable as proof of a difference.
- the degree of confidence with which the level of significance will lead to the correct conclusion.

Unfortunately many of these factors cannot be defined before commencement of a study as, by definition, the parameter that is likely to be affected in screening tests is unknown.

One aspect that is known, however, is that in untreated animals major malformations occur at a very low frequency so that if they were to be the sole criterion of effect then statistical theory would demand the use of hundreds, thousands or even tens of thousands of animals in order to detect small changes in the normal rate (i.e. say doubling of a rate from 0.1% to 0.2%). Experiments of such magnitude are not feasible.

Experience has shown that screening studies involving 15–20 pregnant rodents per group (for each dose) and 3 to 4 groups (three to four doses), can provide a good indication of possible selective effects on the offspring, providing that the studies are well set up and well conducted.

For studies in which two or more methods of examination are employed (i.e. when some litters are delivered by Caesarian section, and other dams are allowed to rear their young, or when some foetuses will be examined solely for visceral anomalies and others for skeletal defects) it would be advisable to allow 10–12 litters for each method of examination.

In performing studies with group sizes of the order of 10–20 animals it is important to consider the studies as initial rather than final steps in the evaluation of safety. Also, one must be aware that effects evoking an all-or-none litter response (e.g. non-pregnancy, abortion) would need to be of the order of 40–50% before they became statistically significant. To detect effects of a lower magnitude the investigator must employ methods of assessment alternative to statistical analysis (e.g. comparison with laboratory background data).

For larger species (cats, dogs, pigs and primates) having a particular value for certain special investigations, the number of animals should be as large as practicable in order to obtain reproducible results.

In these more expensive animals group numbers lower than 10 are conventionally employed. For primary screening there is no scientific justification for this policy and indeed, because of the lack of background information on such species, even larger group sizes would be desirable.

Equivocal results in initial screening tests should be clarified by further experimentation. At this second stage the investigator may have a particular parameter to investigate and therefore can design the study according to correct statistical procedure. Alternatively, the investigator may adjust the dosage and/or dosing period to achieve a more pronounced and recognizable response. It is at this stage of testing that investigators may be able to use the smaller numbers of larger species.

5. Statistical analysis

For teratology studies the basic aim is to determine whether or not a biologically important event has occurred and whether or not it is related to treatment. Undoubtedly the best way to achieve this aim is to design and conduct the study in such a way that the result is self-evident. As experiments are rarely so perfect an investigator may have to resort to additional

aids to assess the meaning of results and plan the next step in the sequence of investigation.

In deciding which method of statistical analysis is the most appropriate the investigator must determine

- the truly independent variables
- their type of distribution.

In teratology studies, doses are usually given to the parent animal and not directly into each individual embryo or foetus; on this basis the litter may represent the unit. The alternative possibility is the analysis based on the individual foetuses, common examples being:

a) Chi² test or Fisher's Exact test on group totals, proportions of live and dead implants, and of abnormal and normal foetuses

b) Analysis of implantation ratios by t-test

c) Analysis of mean pup weight by t- or u-test.

Regarding distributions, litter parameters are usually represented by small discrete variables with only corpora lutea counts and implantation rates approaching the so-called normal (Gaussian) distribution; the distribution of variable young is skewed as well as being composed of discrete intervals whilst the incidence of dead young (resorptions) and abnormal young tends to follow a Poisson distribution. This peculiarity of litter parameters effectively devalues the reliability of statistical analysis such as the Student's 't' test and variance analysis based on the normal curve for continuous variables.

In general terms investigators will find that in most cases, but not all, non-parametric methods will provide the simplest and most effective methods of analysing the results of teratology studies. For example, for all-or-none litter responses such as non-pregnancy, total resorption and/or abortion, Fisher's exact test (or tables of hypergeometric progression) is one of the more useful methods of assessment; when control values are nil or very low a one-tailed criterion should be employed. As an alternative the Chi² test may be used if there are a sufficient number of litters (i.e. not less than 5 in each compartment of a 2×2 contingency table).