# Testing for teratogenicity: the current situation in Europe

Erminio Giavini, Department of Biology, University of Milan, Italy

## Introduction

Awareness of teratogenicity of chemicals increased greatly in the early 1960s when the dysmorphogenic effects of thalidomide (used at that time as antiemetic in pregnancy) primarily on human limb development, was recognized. Since the recognition of prenatal vulnerability of the human conceptus, much has been done to detect potential teratogens and to regulate human exposure to them. It was clear from several experimental studies that chemically induced effects on prenatal development are not limited to morphological abnormalities: pre- and peri-natal death, growth retardation, behavioural and functional impairment are all parts of the spectrum of alterations induced by exposure to chemicals in utero; as a consequence, nowadays the term “developmental toxicity” is preferred to “teratogenicity”.

The process of developmental toxicity testing of environmental chemicals is ruled by a framework of guidelines which have many methodological features in common with those applied for pharmaceuticals testing. There are also some differences due to the characteristics of the two substance classes: while drugs are given to people intentionally and in known doses, environmental chemicals are ubiquitous and both the doses and the exposed population are much more complicated to control. This results in a much higher degree of uncertainty in the case of environmental exposure level and the target population (Buschmann, 2006).

## Testing of environmental chemicals

The most recent guidelines for testing the effects of chemicals on development are quite similar in different countries, with just minimal differences. This is due to the on-going work done by the OECD (Organization for Economical Co-operation and Development) committee. The OECD is an international organization composed of more than 20 industrialized member countries. One of its activities is to develop testing protocols for health and environmental effects that would be jointly acceptable to all its member

countries in order to reduce the necessity of repeating tests to fulfil the requirements of individual countries.

The most relevant OECD guidelines for testing developmental effects of chemicals are the following:

- Prenatal developmental toxicity study (2001)
- One- generation reproduction toxicity study (1983)
- Two-generation reproduction toxicity study (2001)

Very similar guidelines have been adopted by the European Commission. Actually, the EU prenatal developmental toxicity study is a replicate of OECD TG 414 (2001). This test is designed “to provide general information about effects of prenatal exposure on the pregnant animals and on the developing organism in utero”, including assessment of maternal effects as well as death, structural abnormalities, or altered growth in the foetus. So, the test has been designed to discover the intrinsic property of a chemical to produce adverse effects on conceptuses (hazard identification).

In summary, the test substance must be administered to pregnant animals from implantation to one day prior to the day of scheduled kill which should be as close as possible to the normal day of delivery (usually day 21 of gestation in rats and day 29 of gestation in rabbits). It is recommended that testing be performed in one rodents species (preferred rat) and in one non-rodent species (rabbit). The animals are divided in at least four groups containing a sufficient number of females to result in approximately 20 pregnant animals/group. At least three dose levels and a concurrent control should be used. This last group should be a sham-treated control group or a vehicle-control group if a vehicle is used in administering the test substance. It is very important to use different dose levels because embryotoxic effects (i.e all the possible adverse effects on a conceptus) are strictly dose-related and the dose levels should be spaced to produce a gradation of toxic effects. In particular, to maximize the possible adverse effects and to be sure that the conceptus will be exposed to the highest possible level, the highest dose should be chosen with the aim to induce some maternal toxicity (e.g., a minimal decrease in maternal body weight gain in comparison to control). The lowest dose level should not produce any evidence of either maternal or developmental toxicity, it should be the No Observed Adverse Effect Level (NOAEL). It is very important to find out the NOAEL because this dose could be necessary in case of risk assessment evaluation.

The test substance (or vehicle) is usually administered daily, orally by intubation. If other routes of administration are used, they must be justified. In order to verify signs of toxicity

(or abortion) the pregnant animals must be observed after treatment and their conditions recorded; besides, they must be weighed at least every three days during the treatment period and on the day of scheduled kill. Immediately after killing the dams must be necropsied in order to verify pathological changes. The uteri must be removed and carefully examined in order to count the number of implantation and signs of resorptions (dead embryos). The foetuses must be examined for external abnormalities, weighed and afterwards examined for visceral and skeletal abnormalities using adequate methods.

Unlike pharmaceuticals, environmental chemicals should be also tested through a "Reproduction toxicity study". The preferred test for chemicals with a production of more than 1 tonn/year is the "Two generation reproduction toxicity test". This study is designed to provide information concerning the effects of the test substance on the male and female reproductive systems (i.e., general aspects of fertility), gestation, parturition, lactation and weaning, and the growth and development of the offspring. The rat is the preferred species for testing and the test substance should be administered by oral route (mixed in the diet or water, or by intubation). One control and three test groups are required. The males should be treated for at least one complete spermatogenic cycle (about 70 days) before mating; the females for at least three oestrous cycle (2 weeks) before mating, during mating, pregnancy and lactation. At weaning, the treatment is continued to F1 offspring during their growth into adulthood, mating and production of an F2 generation, until the F2 generation is weaned. Although this test is not finalized to verify embryotoxic effects, it may furnish important information on number of live pups, their weight at birth, and postnatal viability and development.

Classification and labelling of chemicals

On the basis of the results of these experimental studies and other information when available (for example, epidemiological studies), substances are categorized and labelled according to Appendix VI to the Commission Directive 93/21/EEC.
As far as developmental toxicity is concerned, substances are divided into three categories (plus the not classified substances):

- Category 1 includes "substances known to cause developmental toxicity in humans" on the basis of a sufficient evidence of a causal relationship between human

exposure to the substance and subsequent developmental toxic effects in the progeny;

- Category 2 includes “substances which should be regarded as if they cause developmental toxicity to humans” on the basis of clear results in appropriate animal studies where developmental toxic effects have been observed at dose levels that have not produced maternal toxicity or at the same dose levels able to produce maternal toxicity but which are not a consequence of maternal toxicity (maternal suffering may have effects on conceptuses);
- Category 3 includes “substances which cause concern for humans owing to possible developmental effects”. Also in this case the classification is based on results obtained in appropriate animal studies (as for category 2), but where the evidence is insufficient to place the substance in category 2.

Consequently, category 1 is based on the results of human epidemiological studies, while categories 2 and 3 are based on the results of experimental studies. Substances may be not classified by two reasons: either available data show that they do not have developmental toxic potential or due to the lack of data.

If a substance is classified into one of the above categories, it must be labelled with the following risk phrases:

- category 1 and 2: R61 may cause harm to the unborn child;
- category 3: R63 possible risk of harm to the unborn child.

It must be stressed that, according to Directive 76/769/EEC, substances classified in category 1 or 2 for reproductive/developmental toxicity may not be used on the market for sale to the general public.

In well conducted developmental toxicity studies it is possible to obtain a NOAEL. This dose level is very important in order to extrapolate the data obtained in animals to humans. In this process safety factors (SF) are applied to obtain an Admissible Daily Intake (ADI) of the substance, i.e. a dose to which the population can be exposed daily without risk for health.

The NOAEL is strictly related to the threshold dose. The question about the existence of a threshold dose in chemical teratogenesis dates more than 25 years ago when some scientists regarded the mechanisms of teratogenesis similar to those inducing cancer on the basis of single cell mutation. On the contrary, teratogenic phenomena were considered

threshold phenomena, not stochastic phenomena by the majority of the teratologists. Stochastic phenomena (some kinds of cancer and mutation induction) are produced by changes in DNA and, therefore, can be the result of changes in a single cell. In these kinds of phenomena the risk decreases with the dose but never theoretically disappears. Other characteristic of stochastic phenomena is that while the risk increases with doses, the magnitude of the induced disease does not. If chemical teratogenesis is a threshold phenomenon, this means that the majority of teratogenic agents have a “no effect dose”. This concept was based on the observation that teratogenesis is a multicellular phenomenon, i.e., a large number of embryonic cells must be damaged in order to obtain a malformation. Furthermore, both the incidence and severity of malformations increase with the dose, a non stochastic characteristic (Brent, 1986; Giavini, 1988). Recently, studies about the mechanisms of action of teratogenic agents evidenced a possible interaction between genes and teratogens in the induction of congenital malformations (Finnell et al., 2002; Giavini and Menegola, 2003). Embryonic development is under the control of several genes regulating the main processes typical of development: cell proliferation, migration, differentiation and death (apoptosis). This control is very complicated and involves the interaction of different genes through transcription factors, i.e. proteins able to modulate the expression of specific genes. Environmental chemicals could be teratogenic by interfering with transcriptional response or with proteins involved in the signal transduction pathways that activate the expression regulation of developmental genes. Recently Menegola et al. (2006a) showed that chemicals able to inhibit histone deacetylase as Valproic acid, Trichostatin A, Apicidin, are teratogenic. The mechanism is the histone hypercethylation of embryonic chromatin with consequent alteration of gene expression. Although these new mechanisms of action imply the involvement of direct or indirect alterations of gene activity, they do not modify the concept of teratogenesis as a threshold phenomenon: also in these cases a threshold number of embryonic cells must be involved to obtain a malformation, which means that a threshold dose does exist.

However, the NOAEL not necessarily corresponds to the threshold dose. This is due to a relatively small number of samples in a developmental toxicity study. As a consequence, the NOAEL is just a statistical approximation of a threshold dose. For this reason, in order to extrapolate the data obtained in animals to humans, safety factors are used.

Safety factor is usually 100, based on the assumption that 10 is an expression of intraspecies variability of response and another 10 is expression of interspecies variability. The ADI is equal to NOAEL/SF: for example, if in one experiment the NOAEL is 50 mg/kg

body weight and a SF of 100 is applied, the ADI will be 0.5 mg/kg body weight, i.e about 30 mg/day for a person of 60 kg. Justified deviations in the value of SF are possible (increase or reduction).

Classification and labelling of chemicals is responsibility of the European Chemical Bureau (ECB) which was established within the Environment Institute of Joint Research Center in Ispra (Italy) since 1993. The ECB is responsible for the Technical Committee on Classification and Labelling of Dangerous Substances (TC C&L). The TC C&L discusses classification and labelling of substances of special concern which are then proposed to be included into the list of harmonised classification of substances, Annex 1 of Directive 67/548/EEC. The Annex 1 contains at present about 2700 existing and 1100 new chemicals. Prioritising of substances to be put into agenda is based on several characteristics of substances, but first of all concern about carcinogenic, mutagenic and reproductive toxicity (including developmental toxicity). Experts from the Authorities of the Member States are invited to participate at the meetings of the TC C&L and they should come to agreement on the recommendations that the ECB subsequently forward to DG ENV. Experts from industry and NGOs may be also invited as observer and can participate to the scientific discussion, but not the final decision. For substances with carcinogenic, mutagenic and reproductive toxicity concern, the TC C&L may ask to consult specialised experts.

On December 18 2006, the European Commission approved a new system for Regulation, Evaluation and Authorisation of Chemicals (REACh, Regulation EC 1907/2006)). This system applies both to existing (marketed before 1981, about 100.000) and new chemicals (marketed after 1981, about 3000). Under the REACh legislation the industries (manufacturers, importers, users) will play a pivotal role in classifying and labelling substances according to the standard criteria of the EU's chemical registration. There will be a general obligation for industries to submit a registration for each substance including a technical dossier containing information on the properties, uses and classification of a substance. The dossier evaluation will be made by Member States Authorities which may require further information. For substances with very high concern (e.g., with carcinogenic, mutagenic or reproductive toxicity properties) authorization is required for their use and marketing. An European Chemical Agency will be created and

located in Helsinky to manage the technical, scientific and administrative aspects of REACh system and to ensure consistency of decision making at Community level.

As far as the pesticides are concern, they are controlled by the Plant Protection Products Directive (91/414/EEC), which came into force on July 26 1993. The Directive provides for the establishment of a positive list of active substances, that will be include in the Annex 1, that have been shown to be without unacceptable risk to people and environment. Before an active substance can be considered for inclusion in Annex 1, companies must submit a complete dossier on the active substance, including toxicity data. The dossier is examined by one Member State (rapporteur) for evaluation. A report of evaluation is submitted to the European Food Safety Authority (EFSA) for a peer review, afterward EFSA makes a recommendation to the European Commission on whether the inclusion in the Annex 1 is acceptable. This recommendation is discussed by all Member States in the framework of the Standing Committee on the Food Chain and Animal Health.

As shown by the cited laws, directives, and evaluation processes, the EEC has made and is making a great effort in order to verify the possible risks for human health (including effects on development) due to the exposure to environmental chemicals.The main problem is related to the “existing chemicals” (more than 100.000 substances) for which possibly very little toxicological information is available.

Critical aspects

The results obtained in experimental studies are not enough to consider a substance as a developmental toxicant for human conceptus. Evidences of developmental toxicity in laboratory experiments can be used as precautionally indicative of possible risks to humans in the case of pharmaceuticals. For the environmental chemicals numerous differences between experimental conditions and human real situation do not permit the direct application of results from animal models in assessing potential human effects. Hazard identification, in this case, is only the first step of the assessment process, not the end in itself. Hazard characterization answers the question of whether adverse effects observed in laboratory animals have relevance to humans.
Exposure assessment is crucial to determining the relevance of experimental data to humans. Human exposure data are important for three aims: 1) to identify potentially

exposed populations; 2) to identify and evaluate route(s) of exposure associated to normal chemical use and handling; 3) to estimate the range of likely exposure and obtain quantitative estimates of exposure associated with each pattern of use. These values can be compared with the calculated ADI of the substance in order to quantify the margin of safety.

Several exposure conditions may, per se, drastically reduce the risk for human population. For example, the route of exposure to sodium perborate for workers is the inhalation, but it is limited by local production of hydrogen peroxide with consequent severe and disturbing irritation at the level of nasal and tracheal mucosae. The EU classification and labelling is mainly based on the hazard (i.e. on the intrinsic properties of a substance) rather than on real exposure scenarios (risk assessment) leading, in some cases, to unjustified overestimation of the real risk or overlabelling. Many factors, other than intrinsic properties or human exposure should be taken into account for a final decision about the risk of developmental toxicity in humans. For example, studies on toxicokinetics may provide very important information about the mechanism of action of a developmental toxicant. Toxicokinetics is the decription of the absorption, metabolism, and excretion of a toxic chemical into and from the body. Toxicokinetic data may provide key elements to understanding species differences in response to developmental toxicants and would be also important in the determination of a developmental toxicity study design. For example, the half life of some chemicals is very different in laboratory animals and in humans: phenytoin 3-5 hours in rat, 10 – 60 hours in man; Valproic acid 0.8 hours in mouse and 12 hours in man, cychlophosphamide 0.2 hours in mouse and 7 hours in man. Many chemicals exhibit shorter half-lives in animals than in man. This means that in laboratory animals a series of narrow and high concentrations peaks are obtained during treatment while in man fluctuations are much less pronounced and blood levels are kept within quite constant levels. Unfortunately there are no data of comparative toxicokinetics for environmental chemicals and just few for drugs during pregnancy. For a rational interspecies comparison of risk assessment it should be essential to know if peak levels or constant plasma concentrations are relevant factors determining the teratogenicity of a chemical. In the case of Valproic acid Nau (1985) was able to demonstrate that peaks more than the total drug exposure are important for exencephaly induction in the mouse. Similar results were obtained with caffeine: a comparison of the toxicokinetics and teratogenicity of this drug following injection of a single massive dose or after multiple treatments of divided doses suggested that the teratogenicity is related to peak

concentrations (Jiritano et al., 1985). On the contrary the teratogenicity of cychlophosphamide is likely related to total drug exposure more than to the peak levels.
The guidelines establish a single administration/day, independently of the pharmacokinetics and metabolic characteristics of the substances. As a consequence, if a substance is rapidly metabolized or excreted, the embryo will be exposed only for few hours/day; on the contrary for substances with very low metabolism and excretion there is the possibility of an embryonic overdosage.
An example of a well conducted risk assessment for repro/developmental toxicity have been recently published by Schulte-Hermann et al. (2006).

Although half a century has passed from the introduction of developmental toxicity tets in the safety chemical evaluation, the methods used to detect morphological abnormalities are still those proposed by Wilson (1965) for visceral organs and by Staples and Schnell (1964) for skeleton. In particular, the method used for the evaluation of skeleton is based on the staining of ossified parts with Alizarin red and diafanization of the soft tissues. As the foetal skeleton at term of pregnancy is only partially ossified, single staining of bone cannot accurately evidence all normal and abnormal structures. In spite of the availability of a double staining method (Kimmel and Trammel, 1981) which permits the observation of both ossified and cartilagineous skeleton, the single stain is universally used in routine teratology tests as it is simpler and cheaper than double stain method. However, it has been clearly demonstrated by several academic studies that the double staining can identify a number of skeletal abnormalities not visible with Alizarin staining alone and can also modify the interpretation of some abnormalities: e.g. the extra cervical or lumbar ribs, considered as variations by the single staining users, are considered as malformations by people using double staining on the basis of characteristics of the corresponding vertebrae (Kimmel and Wilson, 1973; Menegola et al., 2002; Wery et al, 2003). Actually, the possibility to see also the cartilagineous parts of the vertebrae, allows a certain characterization of their structure in relation to their position along the skeletal axis.

An important point is related to the effects of mixture of chemicals on development. Pregnant women are usually exposed contemporaneously to several kinds of environmental chemicals due to air and water pollution, food additives, pesticides and chemicals used in the houses or in working places. No data are available about the effects of chemicals mixtures on embryonic development. It is also very difficult to give theoretical

evaluation of the possible effects of mixtures of different chemicals. However, for specific classes of chemicals with the same mechanism of action it is possible to hypothesize additive effects. For example, the triazole fungicides are able to produce facial and axial skeleton malformations in laboratory animals (including frog embryos) (Menegola et al., 2006b). The mechanism of action of these chemicals is the inhibition of Cyp26 embryonic enzymes which control the catabolism of retinoic acid. As all the triazole-derivatives have this property (and also the imidazole-derivatives), it is very likely that the exposure to more chemicals of this class may have an additive effect. This should be taken into account during the evaluation of the exposure for risk assessment.

The main deficiency in the current guidelines for developmental toxicity studies is related to the difficulties in evidencing functional alterations. Everybody knows that this is a possible effect due to in utero exposure to chemicals, but there are no specific guidelines to approach this important aspect. The two generation reproduction toxicity test permits to assess the postnatal development but only in term of physical development or for effects on fertility in the progeny. Several other effects (on endocrine or immunosystem, for example) require specific tests designed on case-by case basis.

## Conclusions.

The EU is in line with the other industrialized countries for testing chemicals for developmental toxicity effects. The great problem of the enormous quantity of existing chemicals, often without significant toxicological information, will be the aim of the new REACh regulation which, in the first step, will examine about 30,000 chemicals prioritized on the basis of tonnage of production.
In the meantime it is necessary to introduce the concept of “realistic exposure scenario” as one of the principles on which the classification of environmental chemicals is based; besides, the OECD or other organizations should identify new guidelines for specific functional effects (e.g. endocrine disruptors) and the industries should be encouraged to use more advanced methods for finding alterations induced by in utero exposure to environmental chemicals.

## References

Brent R.L.: Editorial comment: definition of a teratgen and the relationship of teratogenicity to carcinogenicity. Teratology, 34: 359-360, 1986.

Buschmann J.: Critical aspects in reproductive and developmental toxicity testing of environmental chemicals. Reprod. Toxicol. 22: 157-163, 2006

Commission Directive 93/21/EEC of 27 April 1993 adapting the technical progress for the 18th time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging, and labelling of dangerous substances. Official Journal of European Communities L110, 1993.

Council Directive 91/414/EEC of 15 July 1991 concerning the placing of plant protection products on the market. Official Journal of European Communities L225, 2001.

Directive 67/548/EEC, general classification and labelling requirements for dangerous substances and preparations, Annex VI. Official Journal of European Communities L225/287,.2001.

European Commission. Directorate General. Joint Research Centre. Institute for Health and Consumer Protection. Testing methods and directives. 2004. Available: http://ecb.jrc.it/DOCUMENTS/Testing-Methods/ListofojwebMay2004.pdf.

Finnell R.H., Gelineau-van Waes J., Eudy J.D., Rosenquist T.H.: Molecular basis of environmentally induced birth defects. Ann. Rev. Pharmacol. Toxicol.42: 181-208, 2002.

Giavini E.: Evaluation of the threshold concept in teratogenicity studies. Teratology, 38: 393-395, 1988.

Giavini E., Menegola E.: Gene-teratogen interactions in chemically induced congenital malformations. Biol. Neonate, 85: 73-81, 2003.

Jiritano L., Bortolotti A., Gaspari F., Bonati M.: Caffeine disposition after oral administration to pregnant rats. Xenobiotica, 15: 1045-1051, 1095.

Kimmel C.A. and Trammel C.: A rapid procedure for routine double staining of cartilage and bone in fetal and adult animals. Stain Technol., 56: 271-273, 1981.

Kimmel C.A. and Wilson J.G.: Skeletal deviations in rats: malformations or variations? Teratology, 8: 309-316, 1973.

Menegola E., Broccia M.L., Di Renzo F., Giavini E.: Comparative study of sodium valproate- induced skeletal malformations using single or double staining method. Reprod. Toxicol., 16: 815-823, 2002.

Menegola E., Di Renzo F., Broccia M.L., Giavini E.: Inhibition of histone deacetylase as a new mechanism of teratogenesis. Birth Defects Res. (Part C) 78: 345-353, 2006a.

Menegola E., Broccia M.L., Di Renzo F. Giavini E.: Postulated pathogenic pathway in triazole fungicide induced dysmorphogenic effects. Reprod. Toxicol., 22: 186-195, 2006b.

Nau H.: Teratogenic valproic acid concentrations: infusion by implanted minipumps vs. conventional injection regimen in the mouse. Toxicol. App. Pharmacol. 80: 243-250, 1985

Nau H.: Species differences in pharmacokinetics and drug teratogenesis. Environ. Health Perspect., 70: 113-129, 1986.

.

OECD. Test guidelines 415. OECD guideline for testing of chemicals. One-generation reproduction toxicity study. 1983.

OECD. Test guideline 414. OECD guideline for testing of chemicals. Prenatal developmental toxicity study. 2001

OECD. Test guideline 416. OECD guideline for testing of chemicals. Two-generation reproduction toxicity study. 2001.

Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive

1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/796/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/ec and 2000/21/EC. Official Journal of the European Union L396, 2006.

Schulte-Hermann R. and other 14.: Analysis of reproductive toxicity and classification of glufosinate-ammonium. Regulatory Toxicol Pharmacol. 44: S1 –S76, 2006.

Staples R.E. and Schnell V.L.: Refinements in rapid technic in the KOH Alizarin red S method for fetal bone. Stain Technol., 39: 61-63, 1964.

Wilson J.G.: Methods of administering agents and detecting malformations in experimental animals. In “Teratology: Principles and Techniques” (Wilson JG and Warkany J eds.), pp. 262-277. The University of Chicago Press, Chicago, 1965.

Wery N., Narotsky M.G., Pacico N., Kavlock R.J., Picard J.J., Gofflot F.: Defects in cervical vertebrae in boric acid-exposed rat embryos are associated with anterior shifts of hox gene expression domain. Birth Defects Research (A) 67: 59.67, 2003.