Reproduction and Development

Horst Spielmann

Bundesinstitut für gesundheitlichen Vebrauchershutz und Veterinärmedizin, Berlin, Germany

The currently developed and validated *in vitro* tests for female and male fertility and also for developmental toxicity are described and evaluated according to their potential use as screening or replacement alternatives to the established *in vivo* tests in reproductive and developmental toxicology. Alternative methods today can only be used to evaluate a few specific components of the integrated reproductive functions in both females and males. However, in the field of developmental toxicity testing there is a strong theoretical and empirical basis for the predictive power of *in vitro* screens using mammalian embryos as well as embryonic cells and tissues. Several of these assays have been validated or are currently undergoing validation in several laboratories and are >80% concordant with *in vivo* results. Failure to achieve 100% accuracy reflects the inherent limitations of these systems, which are manageable, as the concordance rates are still good. The level of concordance suggests that these assays are adequate for screening purposes to complement traditional *in vivo* testing. The use of these assays as screens will save valuable *in vivo* testing resources for those compounds most likely to enter the market and to which people will be exposed. — *Environ Health Perspect* 106(Suppl 2):571–576 (1998). *http://ehpnet1.niehs.nih.gov/docs/1998/Suppl-2/571-576spielmann/abstract.html*

Key words: *in vitro* test, toxicology, reproduction, development, embryo, embryotoxicity, evaluation, validation

Introduction

Reproduction is a continuous cycle. For the purposes of toxicity testing, however, it is broadly divided into pregnancy in females, including prenatal or postnatal developmental toxicity, and the remainder of the cycle in both males and females during which fertility may be impaired.

During the past 20 years research in reproductive toxicology has focused on the use of alternatives to living mammals for testing the potential reproductive toxicities of chemical and physical agents. Recent reviews include an Organisation for Economic Co-operation and Development workshop in Ottawa, Canada, in 1992 (*1*) and a European Centre for the Validation of Alternative Methods (ECVAM) workshop in Ispra, Italy, in 1994 (*2*). International experts concluded at both conferences that the use of *in vitro* methods is well established and that they are invaluable for conducting mechanistic reproductive toxicity studies. Additionally, *in vitro* methods already play a valuable role in so-called secondary testing, i.e., in the screening of series of structurally related chemicals when at least one of the chemicals is of known reproductive toxicity *in vivo*.

The majority of research into the development of alternative tests has concentrated on teratogenicity, which is only one manifestation of adverse effects on development and does not cover fertility, which includes sexual behavior, spermatogenesis, oogenesis, fertilization and the development of the zygote up to term, postnatal development, and hormonal activity.

In Vivo Testing for Regulatory Purposes

Currently, reproduction/development screening tests (*3,4*) or multigeneration studies must be conducted to provide information on the effects of industrial chemicals on all aspects of the highly complex reproductive cycle (*5,6*). For chemicals used as drugs, segment studies must be conducted covering three important phases of pre- and postnatal development and fertility (*7*). Because of the complexity of the reproductive cycle, from gamete maturation through implantation of the early embryo into the uterus, and because of the lack of validated alternative tests for most steps in the cycle, testing in living animals is the only option currently available for assessing the possible effects of chemicals on reproduction. Moreover, because of the complexity of functions that are only found in living animals, *in vitro* screening may never be able to cover all of the aspects of fertility. Thus, the key question is whether sufficient information can be derived from alternative tests to be able to classify and label chemicals as toxic to the reproductive system.

Fertility: *In Vitro* Approaches

Female Fertility

Some aspects of female reproductive function can be modeled *in vitro*, and several cellular components of the female reproductive organs can be maintained in culture (*8*). Although none of the organs have been used or validated as toxicity screens, several may be useful for specialized toxicologic studies. In the future, a battery of such systems may be able to cover a large proportion of the female reproductive cycle.

In females, the proliferation of primordial germ cells and the initial steps of meiosis occur long before birth. From puberty onward, a small number of the primary oocytes complete oogenesis and are released from the ovary. Ovarian somatic cells (granulosa, thecal, and stromal cells) can be maintained in culture (*8*) and any adverse effects can be assessed by examining cell morphology, viability, and hormonal responsiveness.

Techniques for *in vitro* fertilization (including functional maturation of spermatozoa) are routine, both clinically and in farm and laboratory animals. The methods used have not been validated for testing purposes, although some toxicologic studies have been performed (*9*). For example, the mammalian preimplantation period can be investigated by culturing embryos from the first cleavage divisions up to implantation,

This paper was prepared as background for the 13th Meeting of the Scientific Group on Methodologies for the Safety Evaluation of Chemicals (SGOMSEC): Alternative Testing Methodologies held 26–31 January 1997 in Ispra, Italy. Manuscript received at *EHP* 9 May 1997; accepted 29 May 1997.

Address correspondence to Dr. H. Spielmann, Zentralstelle zur Erfassung und Berwetung von Ersatz- und Ergänzungsmethoden zum Tierversuch, BgVV, Diedersdorfer Weg 1, D-12277 Berlin, Germany. Telephone: 49 30 8412 2270. Fax: 49 30 8412 2958. E-mail: zebet@bgvv.de

Abbreviations used: CERC, chick embryo neural retina cells; CHEST, chick embryotoxicity screening test; CNS, central nervous system; ECVAM, European Centre for the Validation of Alternative Methods; ES, embryonic stem; EST, embryonic stem cell test; FETAX, frog embryo teratogenesis assay *Xenopus*; HEPM, human embryonic palate mesenchymal; LIF, leukemia inhibitory factor; MOT, mouse ovarian tumor.

and toxicologic investigations at the chromosomal level have been published (*10*). Thus, there are no alternative tests available in the near future that would enable the screening of chemicals for female reproductive toxicity with the predictivity required for the safety assessment of chemicals. In the long term, a complex battery of *in vitro* tests may be devised.

Male Fertility

There is particular interest in the development of alternative approaches for assessing male reproductive toxicity. The human male has a relatively low sperm count; the number of sperm per ejaculate is typically only between 2- and 4-fold higher than that at which fertility is significantly impaired. In contrast, the number of sperm in a rat or rabbit ejaculate is many times (up to 1000-fold) that which will produce maximum fertility. Epididymal sperm count can be reduced by as much as 90% in the rat without significantly affecting fertility. Consequently, animal models may be insensitive indicators of human reproductive hazards. Studies on male reproductive toxicity are aided by the ready availability of human target cells.

The production of spermatozoa from stem cells is a complex process that takes about 5 weeks in mice and 11 weeks in humans. Chemicals can disturb normal spermatogenesis by direct interaction with targets within the testis itself, or indirectly by interfering with hormonal stimulation or alterations in blood supply.

It is not possible to mimic the whole of the male reproductive cycle *in vitro*, but several components can be studied individually. Although they have been used extensively in toxicologic studies, they still do not represent a viable alternative to *in vivo* tests. Male germ cells are produced from stem cells throughout mature life. Thus, it should be easier to devise culture systems that are able to support the whole of spermatogenesis.

Several testicular cell types can be maintained in culture, either alone or in combination; these include Sertoli–germ cell cocultures, Sertoli cell-enriched cultures, germ cell-enriched cultures, Leydig cell cultures, and Leydig–Sertoli cell cocultures. All of these systems have been used successfully to study specific features of testicular toxicity (*11*). Primary cultures of testicular cells retain many of the *in vivo* characteristics but only have a limited life span. The ability to study individual cell populations from a heterogeneous organ such as the testis is a powerful tool for probing mechanisms of toxicity. However, the loss of interactions with other cell types is a serious limitation to their use for screening purposes.

Additional measurements may be incorporated into current *in vivo* testing protocols. High quality histopathology of the testes and epididymis would enable effects on specific cell populations to be evaluated. Alternatively, recently developed flow cytometric techniques could be used. One of the flow cytometric methods developed for detecting alterations in spermatogenesis allows simultaneous measurement of cellular DNA content, RNA content, and stainability (*12*). This procedure provides a rapid assessment of up to eight different testicular cell populations.

Semen Analyses

There are several techniques available for monitoring sperm motility, morphology, and various other aspects of semen composition, including the fertilizing ability of sperm. Such approaches could be used for both human and animal semen analyses. The direct addition of test chemicals to semen samples *in vitro* may be a valuable approach, given the availability of human material. Fertilizing capacity declines with increasing proportions of abnormal sperm head morphology, and there is a high correlation between chemical-induced sperm head abnormalities and altered sperm chromatin structure in the mouse (*13*).

Developmental Toxicity: *In Vitro* Tests

Over the past 20 years, more than 30 different culture systems have been proposed as tests for developmental toxicity. The majority of these tests have each been used by only one laboratory. The culture systems fall into four categories: established cell lines, primary cell cultures, nonmammalian embryos, and mammalian embryos or primordia.

Cell Lines and Embryonic Stem Cells

Established cell lines that have been used for developmental toxicity screening include human embryonic palate mesenchymal (HEPM) cells (*14*), mouse ovarian tumor (MOT) cells (*15*), and neuroblastoma cells (*16*). The results of a blind trial with a dual HEPM/MOT approach showed an unacceptably high level of false positives (> 50%) (*17*).

In a more recently developed approach, blastocyst-derived totipotent embryonic stem (ES) cell lines of the mouse were used for *in vitro* embryotoxicity testing. ES cells can be maintained in an undifferentiated state in the presence of feeder layers and/or purified leukemia inhibitory factor (LIF). When the LIF is removed, ES cells differentiate into a variety of cell types depending on the culture conditions. For example, in the mouse, muscle cell differentiation from ES cells reflects myogenesis *in vivo* (*18*), and the development of hemopoietic cells parallels hematopoiesis in the developing embryo (*19*).

ES cells offer several new approaches with respect to screening for embryotoxicity *in vitro*, enabling the use of differentiating permanent embryonic cells. Cytotoxicity (*20*) and effects on differentiation (*19,21*) have been used as end points in embryotoxicity tests with ES cells; inhibition of the differentiation of ES cells in a micromass culture has proved a particularly promising assay under routine testing conditions (*21*). Determination of the two essential features of embryotoxic agents (i.e., inhibition of differentiation combined with a higher sensitivity of embryonic cells than adult tissues to cytotoxic damage) in a single assay procedure with ES cells was recently attempted (*22*). The predictive value of this embryonic stem cell test (EST) was as good as the results obtained with *in vitro* embryotoxicity tests using either rodent whole embryos or embryonic tissues.

ES cells are routinely used in the production of transgenic mice, and methods to introduce targeted mutations and reporter constructs are well established. Transgenic markers could be devised to simplify the end points used in a particular toxicity test and to enable the automation of such assays. ECVAM supports the continued development of these approaches.

Aggregate and Micromass Cultures

Aggregates of primary cultures of chick embryo neural retina cells (CERC) provided encouraging results as a screen for developmental toxicity (*23*). The CERC assay may not have gained wider acceptance because it is based on the differentiation of nonmammalian cells from an organ that is not the primary target of embryotoxic agents in the human embryo.

When cells from the undifferentiated mesenchyme of early chick embryo limbs were cultured in small volumes at high density, they formed numerous small foci of differentiating chondrocytes within a background of apparently undifferentiated cells. Cell adhesion, movement, communication,

division, and differentiation all occur in micromass cultures (*24*). In principle, the micromass test is based on detecting the ability of a particular chemical to inhibit the formation of foci. Embryonic limb or central nervous system (CNS) cells (usually midbrain, which form foci of neurons) from chick, mouse, or rat can be used (*25,26*). The technique has subsequently been modified for use with 96-well microtiter plates (*27*). Cells can be exposed either directly in culture or transplacentally prior to culture (*28*).

Several structure–activity studies have shown that the micromass test can distinguish between teratogens and nonteratogens within a particular chemical class, e.g., retinoids (*28*) and triazole antifungal drugs (*29*). In some cases, organ-, species-, and strain-specific toxicity have been modeled in micromass cultures. For example, ethylenethiourea is more toxic to midbrain than to limb cultures (*30*) and is more toxic to the rat CNS than to the mouse CNS (*31*).

Differentiating cells in both midbrain and limb cultures express cytochrome P450 isozymes (*32*) and, at least in the case of limb cells, these are able to metabolize chemicals such as phenytoin and cyclophosphamide to their toxic metabolites. Validation studies using chemicals from a variety of classes indicate that the percentage of teratogens detected with micromass cultures may vary between approximately 60 to 90%, and that the percentage of nonteratogens identified correctly may vary between 89 and 100% (*26,27,33*). It is possible that much of this variation is accounted for by differences in the exact methodology used, and none of these studies are considered definitive.

Thus, micromass cultures represent robust test systems for studying potential teratogens. It was recommended that the micromass method be included in a comparative trial to determine its applicability relative to several other available *in vitro* systems (*2*).

Embryos of Lower Order Species

Numerous tests that use embryos of submammalian vertebrate and invertebrate species for detecting the teratogenic potentials of chemicals have been described. The organisms that have been used include hydra, fish, frogs, crickets, *Drosophila*, brine shrimp, and slime mold. Several of these are currently being used extensively as models for investigating mechanisms of development. Because any stage or component of development is a potential target for toxicants, the existence of species differences is a strong argument in favor of using vertebrate models for predictive screening. However, subvertebrate systems may have applications in ecotoxicologic monitoring.

Of the nonavian vertebrate systems available, only the frog embryo teratogenesis assay *Xenopus* (FETAX) has undergone limited validation using about 40 different substances (*34*). The overall accuracy in predicting teratogenic potential has been claimed to be 79 to 83% (*34*). FETAX is low cost and rapid and uses a species commonly maintained under laboratory conditions. The assay is limited by the aqueous solubility of test substances, the relative lack of validation, and the small number of laboratories that have used the system. Nevertheless, it has been recommended that FETAX be included in a comparative trial of alternative tests for developmental toxicity (*1,2*).

Avian Embryos

Although avian embryos are widely used as models in developmental biology, they have rarely been used for embryotoxicity testing. The chick embryotoxicity screening test (CHEST) was devised by Jelinek and co-workers (*35*) and has been used extensively in their laboratory, but not elsewhere. Intraamniotic injection eliminates the problem of continuous exposure of the embryo because the test substance is readily distributed to the extraembryonic compartments. Growth retardation, malformation, and death as well as dose–response and stage–response relationships and malformation spectra are easily determined. The results obtained from testing over 130 compounds have been published (*35,36*). One general problem with CHEST has been the inability to distinguish general toxicity from specific developmental effects.

Mammalian Whole Embryo Culture

Mammalian embryos can be maintained in culture for short periods throughout the phase from fertilization to the end of organogenesis (*10*). For toxicity testing, the period from the end of gastrulation to midorganogenesis has been investigated extensively. Screening systems using mouse (*37*) and rat (*38*) embryos have been proposed, and the culture of rabbit embryos has recently been optimized (*39*).

Head fold or early somite stage embryos are dissected free from maternal tissue, parietal yolk sac, and Reichert's membrane, leaving the visceral yolk sac and ectoplacental cone intact. The conceptus is cultured in medium under defined gassing conditions for 24 to 48 hr, usually in a roller bottle system. A variety of media have been used, all of which contain a high proportion of serum. Rat serum is most common (*40*), but mouse, rabbit, cow, monkey, and human sera (*41*) have been used. The test compound can be added to the cultures for defined periods or for the entire culture period. Metabolic activation systems can be incorporated, including the addition of S9 or microsomal fractions of liver from different species, coculture with hepatocytes, sequential hepatocyte/whole embryo culture, and the addition of serum from treated animals or humans (*42*).

At the end of the culture period a number of end points can be measured, including effects on the development of the visceral yolk sac vascularization and circulation; effects on hematopoiesis, embryonic growth (e.g., size and protein and DNA contents); and differentiation (number of somites, morphologic score); and dysmorphogenic effects (*37,38,43*). Interpretation of the results takes into account adverse effects on yolk sac development and embryonic growth and differentiation as well as adverse effects specifically on dysmorphogenesis. Validation studies have been carried out (*43*) and an interlaboratory validation study has been conducted (*44*). In a validation study on different culture systems, six pairs of coded compounds were tested in chick and rat embryo cultures and in brain cell aggregate cultures (*45*). Bechter et al. (*46*) reported an excellent agreement between *in vivo* and *in vitro* data for a series of retinoids.

Mammalian whole embryo culture systems are well developed *in vitro* tests for the detection of potentially teratogenic compounds and for the elucidation of mechanisms of teratogenicity. This *in vitro* system has been used in many academic and industrial laboratories and has proved to be a valuable tool. It allows the detection of dysmorphogenesis in many organs and the comparison of specific dysmorphogenic effects with general adverse effects on growth and differentiation. In addition, it enables the potencies of structurally related compounds to be ranked. Concentrations of test compounds and metabolites can easily be monitored in the culture medium and embryonic tissues.

However, the system has clear limitations: it is relatively complex, covers only a part of organogenesis, and requires high technical skills. The test can be costly and

it uses mammalian tissue and serum. Whether this is justified with respect to its use as a screen may be evaluated by including it in a comparative trial with other simpler *in vitro* systems.

Toxicokinetics and Metabolism

The production of a direct effect on the developing organism depends on the concentration/time relationship of the chemical and/or its active metabolite(s) in the target cells. Therefore, toxicokinetic and metabolism studies are of crucial importance for the design and interpretation of developmental toxicity studies with both *in vitro* and *in vivo* methods (*47,48*). *In vivo* target concentrations are dependent on maternal absorption of the compound, its distribution, metabolism, and excretion, and its placental transfer and distribution in the embryo. Toxicokinetic studies are also important *in vitro*. The presence of the compound and its stability in the culture medium must be verified, along with an assessment of its transport to, and uptake by, the tissues and cells in culture, its metabolic activation, and its cellular distribution.

Toxicokinetic studies are essential for interpreting results obtained *in vitro* and for extrapolating these to the *in vivo* situation. Activities of added metabolizing systems, such as liver homogenate fractions, isolated enzymes, and hepatocytes, can be assessed by analytical techniques. Measurement of the compound in the cultured tissues and cells is critical, so the target concentration needed to yield an effect can be determined. Such measurements are especially important if little or no activity is observed *in vitro* so that false negatives can be excluded.

Toxicokinetic parameters often differ drastically between *in vivo* and *in vitro* situations. For example, *in vivo* drug levels can fluctuate markedly between doses because of the short half-lives of many chemicals, and high concentration peaks alternate with low or negligible drug levels. In contrast, the chemical is added *in vitro* to the culture medium and may persist for extended periods of time unless it is degraded by hydrolysis or enzymes present in the culture medium.

Discussion

The current status of *in vitro* tests for reproductive and developmental toxicity testing is summarized in Table 1. It is obvious that for reproductive toxicity, alternative methods can only be used to evaluate a few components of integrated reproductive functions both in the female and the male. Table 1 also shows that the situation is more promising in the field of *in vitro* embryotoxicity testing, as many tests can be used for screening purposes and a few have undergone interlaboratory validation with coded chemicals. Some of the tests using mammalian cells and embryos have provided promising results and may be used as screens to set priorities for *in vivo* testing for regulatory purposes, e.g., the ESTs, micromass cultures, and whole embryo cultures. Therefore, ECVAM is currently funding a combined prevalidation and validation trial of the three latter tests according to ECVAM's schemes for the prevalidation and validation of toxicity tests (*49,50*).

One important limitation of the *in vitro* screening tests in developmental toxicity shown in Table 1 is that they have been selected to detect structural alterations but not any other potential manifestations of developmental toxicity. Therefore, one would not expect these tests to predict decreased fetal weight or mortality. Moreover, none of the tests would predict functional abnormalities induced by chemicals.

Therefore, even the best *in vitro* assays are limited in their biology because of the following facts: *a*) none of the tests represent the entire spectrum of developmental events and embryotoxic mechanisms; *b*) the metabolizing capacity of embryonic cells, tissues, and organs is limited; and *c*) the pharmacokinetic parameters *in vitro* are not identical to the situation *in vivo*.

However, the limitations of *in vitro* screens for embryotoxicity can be managed by test selection and design. *In vitro* systems can include specific metabolizing systems or be manipulated to approximate the pharmacokinetic behavior of the test agent in the species of interest. In fact, the ease with which such test systems can be manipulated may in some cases mean that limitations can be turned into advantages.

The primary reason for the general bias of toxicologists against *in vitro* screens is the unrealistic standards that they are

Table 1. *In vitro* tests for reproductive and developmental toxicity.

Test system	End point(s)	Application	Reference
Fertility, rodent and human tissues			
Female			
Ovarian cells: mouse, human	Morphology, hormonal response	Mechanistic studies	(*8*)
In vitro fertilization: mouse, human	Differentiation	Toxicologic studies	(*9*)
Preimplantation embryos	Differentiation	Toxicologic studies	(*10*)
Male			
Testicular cells: mouse, human	Morphology, differentiation	Mechanistic studies Toxicologic studies	(*11,12*) (*11*)
Sperm: mouse, human	Morphology, *in vitro* fertilization	Mechanistic studies	(*13*) (*13*)
In vitro developmental toxicity			
Cell lines: rodent and human			
Ovary tumor cells[a]: mouse	Attachment	Toxicity screening	(*15*)
Palatal cells[a]: human embryo	Proliferation	Toxicity screening	(*12*)
Neuroblastoma: mouse	Differentiation	Toxicity screening	(*16*)
Embryonic stem cell[b]: mouse	Differentiation	Toxicity screening	(*18–22*)
Primary cells and organ cultures: chick and rodent tissues			
Embryonic retina: chick	Aggregation culture	Toxicity screening	(*23*)
Brain aggregates: rat	Aggregation culture	Toxicity screening	(*45*)
Neural tissue[a,b]: rat	Differentiation, micromass	Toxicity screening	(*25,28,30*)
Limb bud cells[b,c]: mouse, rat	Differentiation, micromass	Toxicity screening	(*26,30,31*)
Whole embryo culture: lower species			
FETAX test[a]: frog	Differentiation	Toxicity screening	(*34*)
CHEST test[a]: chick	Differentiation	Toxicity screening	(*35,36*)
Chick embryos	Differentiation	Toxicity screening	(*45*)
Whole embryo culture: mammalian			
Preimplantation embryo: mouse, rat, rabbit	Development	Mechanistic studies	(*10*)
Postimplantation embryo: mouse, rat, rabbit	Differentiation	Mechanistic studies	(*37–39*)
Postimplantation embryo[b,c]: mouse, rat	Differentiation	Toxicity screening	(*40–46*)

[a]Test validated under blind conditions: results ambiguous. [b]Undergoing formal validation in an ECVAM project in 1997/1998. [c]Test validated under blind conditions: result positive.

expected to meet. Despite protests to the contrary, it appears that the *in vitro* tests are expected to be as predictive as, or more predictive than, traditional *in vivo* screens. In other words, they are expected to meet the standards of replacement. For example, it is usually anticipated that *in vitro* tests are able to identify thalidomide as a teratogen despite the fact that thalidomide cannot be identified when tested *in vivo* in the most common rodent species.

Today *in vitro* assays for developmental toxicants can be used for a number of product development and industrial purposes. The three major applications are as follows (*51*): *a*) in the earliest stage of product development, to select from among a group of candidate compounds for a particular indication those compounds that are the least likely to cause developmental toxicity; *b*) to compare the developmental toxicity potential of a new chemical that is only a slight modification of an existing chemical that has already been tested *in vivo*; and *c*) to evaluate compounds for which testing is not routinely performed, usually because the anticipated exposure is very low.

Eventually *in vivo* screening will only have to be conducted on compounds that passed the initial *in vitro* screen. In the latter two instances it may not be necessary to carry out any additional testing *in vivo*.

There are other applications of *in vitro* screens that should be considered seriously. The most significant potential application would be to set priorities for definitive testing of compounds that have been on the market for many years and for which no developmental toxicity data exist to date. Among these untested existing chemicals, *in vitro* screens would permit identifying compounds that may have developmental toxic potential. Using them will save time and money, but they will not replace *in vivo* testing for compounds to which people may substantially be exposed. They will, however, provide information on developmental toxicity for entire classes of compounds for which these data are never routinely obtained. Therefore, using validated *in vitro* tests will increase flexibility in product development and testing without compromising safety. *In vivo* tests will still be conducted on all new pharmaceutical agents and commercial chemicals for which exposure is significant. However, the *in vitro* tests will obviate the need to test *in vivo* materials when there is a high likelihood that their developmental toxicity potential would prevent their introduction into the market.

REFERENCES AND NOTES

1. Villeneuve DC, Koeter HBWM, eds. Proceedings of the International Workshop on *In Vitro* Methods in Reproductive Toxicology. Reprod Toxicol 7:1–175 (1993).
2. Brown NA, Spielmann H, Bechter R, Flint OP, Freeman SJ, Jelinek RJ, Koch E, Nau H, Newall DR, Palmer AK et al. Screening chemicals for reproductive toxicity: the current alternatives. The report and recommendations of and ECVAM/ECTS workshop (ECVAM Workshop 12). ATLA 23:868–882 (1995).
3. OECD. OECD Guidelines for Testing of Chemicals, Guideline 421: Reproduction/Developmental Toxicity Screening Test. Paris:Organisation for Economic Co-operation and Development, 1995.
4. OECD. OECD Guidelines for Testing of Chemicals, Guideline 422: Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test. Paris:Organisation for Economic Co-operation and Development, 1996.
5. OECD. OECD Guidelines for Testing of Chemicals, Guideline 415: One-Generation Reproduction Toxicity Study. Paris:Organisation for Economic Co-operation and Development, 1983.
6. OECD. OECD Guidelines for Testing of Chemicals, Guideline 416: Two-Generation Reproduction Toxicity Study. Paris:Organisation for Economic Co-operation and Development, 1983.
7. International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use. ICH harmonized tripartite guidelines—detection of toxicity to reproduction for medicinal products. In: Proceedings of the Second International Conference on Harmonization, Orlando 1993 (D'Arcy PF, Harron DWG, eds). Antrim, U.K.:Greystone Books, 1994;557–577.
8. Mattison DR. Sites of female reproductive vulnerability: implications for testing and risk assessment. Reprod Toxicol 7:53–62 (1993).
9. Holloway AJ, Moore HDM, Foster PMD. The use of *in vitro* fertilisation to detect reductions in fertility of male rats exposed to 1,3-dinitrobenzene. Fundam Appl Toxicol 14:113–122 (1990).
10. Spielmann H, Vogel R. Unique role of studies on preimplantation embryos to understand mechanisms of embryotoxicity in early pregnancy. Crit Rev Toxicol 20:51–64 (1989).
11. Morrissey RE, Schwetz BA, Lamb JC, Ross MD, Teague JL, Morris RW. Evaluation of rodent sperm, vaginal cytology and reproductive organ weight from NTP 13-week studies. Fundam Appl Toxicol 11:343–358 (1988).
12. Evenson DP. Flow cytometry of acridine orange stained sperm is a rapid and practical method for monitoring occupational exposure to genotoxicants. In: Monitoring of Occupational Genotoxicants (Sorsa M, Norppaed H, eds). New York:Alan R. Liss, 1986;121–132.
13. Evenson DP, Higgins PH, Grueneberg D, Ballachey DE. Flow cytometric analysis of mouse spermatogenic function following exposure to ethyinitrosourea. Cytometry 6:238–253 (1985).
14. Pratt RM, Grove R.I, Willis WD. Prescreening for environmental teratogens using cultured mesenchymal cells from the human embryonic palate. Teratog Carcinog Mutagen 2:313–318 (1982).
15. Braun AG, Nichinson BB, Horowicz PB. Inhibition of tumour cell attachment to concanavalin A-coated surfaces as an assay for teratogenic agents: approaches to validation. Teratog Carcinog Mutagen 2:343–354 (1982).
16. Mummery CL, van der Brink CE, van der Saag PT, De Laat SW. A short term screening test for teratogens using differentiating neuroblastoma cells *in vitro*. Teratology 29:271–279 (1984).
17. Steele VE, Morrissey RE, Elmore EL, Gurganus-Rocha D, Wilkinson BP, Curren RD, Schmetter BS, Louie AT, Lamb JC, Yang LIL. Evaluation of two *in vitro* assays to screen for potential development toxicants. Fundam Appl Toxicol 11:673–684 (1988).
18. Rohwedel J, Maltsev V, Bober E, Arnold HH, Hescheler J, Wobus AM. Muscle cell differentiation of embryonic stem cells reflects myogenesis *in vivo*: developmentally regulated expression of myogenic determination genes and functional expression of ionic currents. Dev Biol 164:87–101 (1994).
19. Heuer J, Graeber IM, Pohl I, Spielmann H. An *in vitro* embryotoxicity assay using the differentiation of embryonic mouse stem cells into haematopoietic cells. Toxicol In Vitro 8:585–587 (1994).
20. Laschinski G, Vogel R, Spielmann H. Cytotoxicity test using blastocyst-derived euploid embryonal stem cells: a new approach to *in vitro* teratogenesis screening. Reprod Toxicol 5:57–64 (1991).

21. Newall DR, Beedies KE. The stem-cell test—a novel *in vitro* assay for teratogenic potential. Toxicol In Vitro 8:697–701 (1994).
22. Speilmann H, Pohl I, Döring B, Moldenhauer F. The embryonic stem cell test (EST): an *in vitro* embryotoxicity test using two permanent mouse cell lines: 3T3 fibroblasts and embryonic stem cells. In Vitro Toxicol 10:199–127 (1997).
23. Daston GP, Baines D, Yonker JE. Chick embryo retinal cell cultures as a screen for developmental toxicity. Toxicol Appl Pharmacol 109:352–366 (1991).
24. Umansky R. The effect of cell population density on the developmental fate of reaggregating mouse limb bud mesenchyme. Dev Biol 13:31–56 (1966).
25. Flint OP. A micromass culture method for rat embryonic neural cells. J Cell Sci 61:247–262 (1983).
26. Flint OP, Orton TC. An *in vitro* assay for teratogens with cultures of rat embryo midbrain and limb cells. Toxicol Appl Pharmacol 76:383–395 (1984).
27. Flint OP. *In vitro* tests for teratogens: desirable endpoints, test batteries and current status of the micromass teratogen test. Reprod Toxicol 7(Suppl 1):103–111 (1993).
28. Kistler A, Tsuchiya T, Tsuchiya M, Klaus M. Teratogenicity of arotenoids (retinoids) *in vivo* and *in vitro*. Arch Toxicol 64:616–622 (1990).
29. Flint OP, Boyle FT. Structure-teratogenicity relationships among antifungal triazoles. In: Handbook of Experimental Pharmacology. Vol 96: Chemotherapy of Fungal Disease (Ryley JF, ed). Berlin:Springer Verlag, 1986;231–249.
30. Tsuchiya T, Takahashi A, Asada S, Takakubo F, Ohsumi-Yamashita N, Eto K. Comparative study of embryotoxic action of ethylenethiourea in rat whole embryo and embryonic cell culture. Teratology 43:319–324 (1991).
31. Tsuchiya T, Nakamura A, Lio T, Takahashi A. Species differences between rats and mice in the teratogenic action of ethylenethiourea: *in vivo*/*in vitro* tests and teratogenic activity of sera using an embryonic cell differentiation system. Toxicol Appl Pharmacol 109:1–6 (1991).
32. Brown LP, Foster JR, Orton TC, Flint OP, Gibson GG. Inducibility and functionality of rat embryonic/foetal cytochrome P450: a study of differentiating limb-bud and midbrain cells *in vitro*. Toxicol In Vitro 3:253–260 (1989).
33. Koelman HJS, Jongeling AJ, van Erp YHM, Weterings PJM, Koopmans ME, Joosten HF, van der Dobbelsteen DJ, van der Aa EM, Yih TD. International ring validation of the *in vitro* micromass teratogenicity test: preliminary results of three Dutch laboratories [Abstract]. Teratology 44:30A (1991).
34. Bantle JA, Fort DJ, Rayburn JR, DeYoung DJ, Bush SJ. Further validation of FETAX: evaluation of the developmental toxicity of five known mammalian teratogens and non-teratogens. Drug Chem Toxicol 13:267–282 (1990).
35. Jelinek R. The chick embryotoxicity screening test (CHEST). In: Methods in Prenatal Toxicology (Neubert D, Merker HJ, Kwasigroch TE, eds). Stuttgart:G. Thieme, 1977;381–386.
36. Jelinek R, Peterka M, Rychter Z. Chick embryotoxicity screening test—130 substances tested. Ind J Exp Biol 23:588–595 (1985).
37. Sadler TW, Horton WE, Warner CW. Whole embryo culture: a screening technique for teratogens? Teratog Carcinog Mutagen 2:243–253 (1982).
38. Schmid BP. Teratogenicity testing of new drugs with the postimplantation embryo culture system. In: Concepts in Toxicology. Vol 3 (Homburger F, ed). Basel, Switzerland:Karger, 1985;46–57.
39. Ninomiya H, Kishida K, Ohno Y, Tsurumi K, Eto K. Effects of trypan blue on rat and rabbit embryos cultured *in vitro*. Toxicol In Vitro 7:707–717 (1993).
40. Bechter R, Teriouw GDC, Lee QP, Juchau MR. Effects of QA 208-199 and its metabolise 209-668 on embryonic development *in vitro* after microinjection into the exocoeiomic space or into the aminiotic cavity of cultured rat conceptuses. Teratog Carcinog Mutagen 11:185–194 (1991).
41. Klein NW, Plenefisch JD, Carey SW, Fredrickson WT, Sackett GP, Burbacher TM, Parker RM. Serum from monkeys with histories of fetal wastage causes abnormalities in cultures of whole rat embryos. Science 215:66–69 (1982)
42. Zhao J, Krafft N, Teriouw GDC, Bechter R. A model combining the whole embryo culture with human liver S9 fraction for human teratogenic prediction. Toxicol In Vitro 7:827–831 (1993).
43. Schmid BP, Trippmacher A, Bianchi A. Validation of whole embryo culture method for *in vitro* teratogenicity testing. In: Developments in the Science and Practice of Toxicology (Hayes AW, Schnell RC, Miyaed TS, eds). Amsterdam:Elsevier, 1983;563–566.
44. Piersma AH, Attenon P, Bechter R, Govers MJAP, Krafft N, Schmid BP, Stadler J, Verhoef A, Verseil C. Interlaboratory evaluation of embryotoxicity in the postimplantation rat embryo culture. Reprod Toxicol 9:275–280 (1995).
45. Kucera P, Cano E, Honegger P, Schilter B, Zijstra JA, Schmid B. Validation of whole chick embryo cultures, whole rat embryo cultures and aggregating embryonic brain cell cultures using six pairs of coded compounds. Toxicol In Vitro 7:785–798 (1993).
46. Bechter R, Teriouw GDC, Tsuchiya M, Tsuchiya T, Kistler A. Teratogenicity of arotinoids (retinoids) in the rat whole embryo culture. Arch Toxicol 66:193–197 (1992).
47. Nau H. Pharmacokinetic aspects of *in vitro* teratogenicity studies: comparison to *in vivo*. In: *In Vitro* Methods in Developmental Toxicology: Use in Defining Mechanisms and Risk Parameters (Kimmel GL, Kochhar DM, eds). Boca Raton, FL:CRC Press, 1990;29–43.
48. Andrews JE, Ebron-McCoy M, Bojic U, Nau H, Kaviock RJ. Validation of an *in vitro* teratology system using chiral substances: stereoselective teratogenicity of 4-yn-valproic acid in cultured mouse embryos. Toxicol Appl Pharmacol 132:310–316 (1995).
49. Curren RD, Southee JA, Spielmann H, Liebsch M, Fentem J, Balls M. The role of prevalidation in the development, validation and acceptance of alternative methods. ATLA 23:211–217 (1995).
50. Balls M, Blaauboer BJ, Fentem J, Bruner L, Combes RD, Ekwal B, Fiedler RJ, Guillouzo A, Lewis RW, Lovell DP et al. Practical aspects of the validation of toxicity test procedures. The report and recommendations of ECVAM Workshop 5. ATLA 23:129–147 (1995).
51. Daston GP. The theoretical and empirical case for *in vitro* developmental toxicity screens and potential applications. Teratology 53:339–344 (1996).