Ann. Ist. Super. Sanità, vol. 29, n. 1 (1993), pp. 97-104

What is a human teratogen: clinical and epidemiological criteria

Roberto BERTOLLINI *(a)*, Massimo PAGANO *(b)* and Pierpaolo MASTROIACOVO *(c)*

(a) OMS, Centro Europeo Ambiente e Salute, Rome, Italy
(b) Centro Internazionale sugli Esiti Sfavorevoli della Riproduzione ICARO-ASM, Rome, Italy
(c) Servizio di Epidemiologia e Clinica dei Difetti Congeniti, Istituto di Pediatria, Policlinico Universitario A. Gemelli, Rome, Italy

Summary. - Known human teratogens explain only 6% of all birth defects. The epidemiological approaches used to study birth defects in human populations were reviewed together with some of the most important methodological problems encountered in this field. The criteria of causality to be met to conclude on the teratogenicity of a given substance were also discussed. A list of the known teratogens with mention of the main birth defects attributed to each of them is enclosed.

Key words: epidemiologic methods, teratogens, malformations.

Riassunto *(Che cosa è un teratogeno: criteri clinici ed epidemiologici)*. - I teratogeni noti fino ad oggi spiegano il 6% di tutti i difetti congeniti. In questa revisione vengono analizzate le metodologie di studio dei teratogeni nell'uomo e i problemi metodologici più importanti che devono essere affrontati per pervenire alla loro identificazione. Vengono anche discussi i criteri necessari per definire che una sostanza o un agente possono causare difetti congeniti. Infine viene fornita una lista completa dei teratogeni conosciuti attualmente.

Parole chiave: metodi epidemiologici, teratogeni, malformazioni.

Introduction

A human teratogen has been defined as a "Chemical, drug, metabolic state, physical agent or psychological alteration during development that has been demonstrated to produce a permanent pathologic or pathophysiologic alteration in the offspring at exposures or circumstances that commonly occur" [1]. The identification of new teratogens is a complex exercise which implies the concurrent analysis of scientific evidence coming from studies carried out using different approaches. In the present paper the principal methodologies used and the problems encountered in studying the teratogens will be reviewed with special attention to the studies in human population groups.

Frequency and causes of congenital abnormalities

The frequency of birth defects among human conceptions is not known. Among spontaneous abortions occurring in recognised pregnancies in the first weeks of gestation (about 25% of all conceptions [2]) the proportion of chromosomal abnormalities, mainly trisomies, varies between 30.5% to 54.9% [3]. Other birth defects have this characteristic (e.g.: neural tube defects) although the proportion among spontaneous abortion is smaller or unknown. This means that when studying a specific malformation recognized at birth and a suspected teratogen, one is missing all the birth defects occurring early in the course of gestation. If a teratogen causes a malformation which is in some cases aborted, it may be difficult to identify it limiting the study to affected newborns.

The frequency of structural congenital abnormalities at birth (also called incidence at birth or birth prevalence) is comparable in several populations and it is around 2% of all births [4]. When abnormalities diagnosed later in life are also considered (mild congenital heart defects, mental retardation, hearing defects, etc.) the rate rises to 4% [5]; a higher frequency up to 15% may be recorded if mild defects are considered.

The distribution of the recognized causes of congenital malformations is summarized in Table 1 [6]. In general it may be estimated that only about 6% of all malformations diagnosed at birth are due to known and preventable teratogenic agents.

Identification of a human teratogen

As for any other cause of disease, the identification of a human teratogen requires the concurrent interpretation of studies of different kind.

Animal studies

Experimental animal studies are the basis for any preliminary evaluation of the reproductive toxicity of a potential teratogen. However the results of these studies cannot automatically be transferred to humans. The metabolic pathways may be different in various species; the disease causing the use of drug as well as other personal and environmental factors (work, smoking habits, drinking habits, nutrition, etc.) may interact with the chemicals and cause a totally unexpected effect. A famous example of the lack of consistency between the results of animal experiments and the human effects is given by the thalidomide. This is a well known cause of limb reduction defects in humans which caused an important epidemic of this birth defect in the sixties [7]: the drug is not teratogenic in most animal experiments with the exception of primates.

Case reports

Human teratogens were often first identified or suspected through spontaneous case reports [8]. In the hands of an alert clinician, case reports may be very powerful instruments to generate hypotheses to be tested by more formal scientific approaches. In case reports one or more babies with a well defined spectrum of defects and an accurately known exposure (in terms of time and quality of exposure) is described. The evidence becomes stronger when is supported by animal experiments, is reported more than once and it is biologically plausible. Most case reports are "false positive": the first observation is not supported by further studies and it is therefore due to chance occurrence. Confirmation of case reports is always needed. Because of the baseline occurrence of birth defects, as described before, it is always possible that a frequently used drug is associated to birth defects solely for the effect of chance. This was, for example, the case of Bendectin, an antinausea drug containing an antihistamine and other compunds; it was associated to birth defects in repeated case reports and withdrawn from the market by the producer after several legal claims. There were animal studies showing the teratogenicity of the drug in animal experiments [9] and case reports claiming the association between the use of the drug and various birth defects. Large well designed epidemiologic studies did not confirm any association between the drug and the disease and attributed the previous results to chance [10].

Table 1. - Percent distribution of the causes of congenital malformations

Cause	Percentage
Genetic anomalies	
Chromosomal anomalies	13.5
Monogenic anomalies	6.0
Multifactorial (interaction of genetic and environmental factors)	20-25
Teratogenic agents	6.0
Maternal diseases	
Epilepsia +/- anticonvulsants	1.5
Insulin dependent diabetes	1.5
Others	<1.0
Prenatal infections (Rubella, CMV, Toxoplasmosis)	2.0
Others (drugs, ionizing radiations alcohol, etc.)	<1.0

Epidemiologic studies

These studies provide quantitative estimates in humans of the presence and strength of the association between a given exposure in pregnancy and the presence of abnormalities in the newborn. Three study designs are mainly used: case control design, cohort design and clinical trials. The main features of each of these approaches are described below.

Case control studies. - In this type of approach, the frequency of exposure to a given agent in pregnancy among cases with a specified malformation (or a pattern of them) is compared with the frequency of exposure of the same agent in a group of infants (controls) without the malformation of interest. When the exposure is more frequent (with a given level of statistical probability) among cases than among controls, an association between exposure and malformation is first established.

This approach is particularly suited for rare diseases (as is the case of most birth defects) and it is attractive for the relative low cost and short time needed to finalize it [11]. Questionnaires or medical records are the most common source of exposure information in case control studies.The major problem is the difficulty to carefully reconstruct exposure history, particularly in relationship with timing of exposure. It has been often claimed that mothers with malformed infants may recall better or differently what happened during pregnancy as compared to women with healthy babies [12]. For this reason the selection of controls among babies with malformations other than the one under study has been suggested [13].

Cohort studies. - In a cohort study a group of pregnant women exposed to a suspected teratogen is followed up to delivery to evaluate the outcome of pregnancy. A

similar not exposed group of women is followed up for comparison. The rate of disease in the two groups is compared and the statistically significance of the difference tested.

This approach allows for a better definition of exposure and a more accurate timing of it. However, because of the rarity of the birth defects, particularly when only specific types are considered, this approach may be impractical, difficult to implement and quite expensive. Large cohorts must be followed up to get a sizeable number of birth defects to study.

Clinical trials. - Clinical trials are similar to cohort studies: a group of pregnant women exposed to a given agent is followed up to evaluate the outcome of pregnancy and compared to a group of unexposed women. In clinical trials exposures (or treatments) are not those occurring "naturally" as in cohort studies but are attributed by the researcher who selects randomly women to be treated and those to receive a control treatment. The random selection precess gives the maximum likelihood of comparability between the two groups for possible factors associated with the effect of the exposure of interest.

Clinical trials are in most of the cases used to evaluate the effectiveness of drugs. In reproductive medicine their use is limited by ethical concerns: it is not possible to prescribe to a pregnant woman a treatment potentially dangerous for the embryo. For this reason clinical trials in this area has been proposed and carry out mainly to test the effectiveness of preventive measures such as, for example, the use of vitamins in the periconceptional period to prevent neural tube defects [14].

Epidemiologic surveillance

In many countries of the world Birth Defects Monitoring Systems have been set up in the past decades to survey the frequency of congenital malformation at birth. An active exchange of information among these registries is routinely carried out to allow for international comparison of birth defect occurrence [3]. The main purpose of the birth defects surveillance is to allow for a quick detection of epidemics of congenital abnormalities and for a rapid identification of the causative agent. The establishment of these registries took place after the thalidomide epidemic, in view of the enormous number of new chemical agents introduced in the environment in the past decades, most of them are not adequately studied for their fetal toxicity.

There are complex methodological problems in the interpretation of data and in the running of these monitoring systems: their intrinsic capability to detect new teratogens has been recently discussed [15].

Methodological issues

Measuring the association

When an association between the increased frequency of a given birth defect and a specific exposure is observed, two are the main objectives of the research: to verify whether the association is real or spurious and to quantify the risk.

To quantify the risk, the measures of association most commonly used are the Relative Risk (RR) and the Attributable Risk (AR). The former is the ratio between the rate of disease among exposed embryos and the rate among unexposed. The latter is the difference between the rates of exposed and unexposed.

The RR measures the strength of the association between a given exposure and the disease without considering the extent of the exposure in the population. In other words a factor strongly associated with a disease can be so rare in the population to have only marginal effects on the total risk of disease. The RR is a number comprised between 0 and infinity; 1 is the value expected in the absence of any association: for example, a RR of 2 indicates that the exposed subjects have a risk of disease two times higher than the unexposed subjects.

The AR conveys an information more relevant for public health purposes than for etiological research. It indicates how much of the disease rate in the population is actually due to the exposure under study. A frequent exposure associated with a relatively low RR may have a high AR and a relevant public health impact.

Table 2. - Latest sensitive period of selected malformations expressed in weeks or days after the last menstrual period (gestational age) [23]

Malformation	Latest sensitive period of susceptibility
Anencephaly	40 days
Spina bifida	42 days
Oesophageal atresia	44 days
Transposition of the great vessels	48 days
Cleft lip	50 days
Diaphragmatic hernia	8 weeks
Anal atresia	8 weeks
Ventricular septal defects	8 weeks
Cleft palate	10-11 weeks
Omphalocele	12 weeks
Hypospadias	14 weeks

Table 3. - Human teratogen list

Drugs	
ACE inhibitors	Intrauterine growth retardation, fetal death, oligoidramnios, neonatal anuria, hypoplasia calvaria
Acid non steroidal Antinflammatory Analgesics (NSAIDs)	Premature closure ductus arteriosus
Androgen hormones	Masculinization of female fetuses
Antiepileptics (AED)	
Phenytoin	Phenytoin or AED syndrome, distal phalanges hypoplasia
Valproic acid	Spina bifida, AED syndrome (?)
Carbamazepine	Intrauterine growth retardation, spina bifida
Trimethadion	Specific syndrome
Antineoplastics	
Folic acid antagonists (Aminopterin, Methotrexate)	Spontaneus abortion, fetal death, specific syndrome
Antimetabolites (Azauridine, Cytarabine, 5-Fluorouracil, 6-Mercaptopurine)	Limb defects, renal defects, central nervous system defects
Alchilant agents (Busulfan, Clorambucil, Ciclofosfamide, Mecloretamine)	Limbs defects, renal defects, central nervous system defects
Antithyroids	
Iodide, I^{131}	Goiter with hypo- or hyperthyroidism
Methimazole	Ulcerlike midline scalp defects
Diethylstilbestrol	Vaginal adenocarcinoma, vaginal adenosis, utero-vaginal defects, female infertility, testicular defects, male infertility
Lithium carbonate	Congenital heart defects, Ebstein's anomaly
Penicillamine	Cutis laxa
Retinoids	
Isotretinoin, Etretinate	A-microtia, hydrocephalus, encephalocele, mental retardation, conotruncal heart defects
Streptomycine, Dihydrostreptomycine, Kanamicine	Hearing deficit
Thalidomide	Phocomelia and specific syndrome
Tetraciclines	Dental staining
Warfarin, Coumarin derivatives	Nasal hypoplasia, condrodisplasia punctata
Unnecessary chemicals	
Alcohol	Feto-alcohol syndrome
Cocaine	Abruptio placentae, disruptive defects

Table 3. - (continued)

Other chemicals	
Biphenyl polychlorurati	Skin and gums hyperpigmentation, conjiunctivitis, icterus
Methyl mercury	Minamata disease
Lead	Mental retardation (mied)
Physical agents	
Cigarette smoke	Intrauterine growth retardation
Ionizing radiations (high doses, at least > 5 rad)	Central nervous system defects, microcephaly, skeletal defects, mental retardation
Biological agents (embryofetal infections)	
Rubella	Cataracts, sensorineural deafness, congenital heart disease (Rubeolic syndrome), intrauterine growth retardation, retinopathy, panencephalitis, endocrinopathies
Cytomegalovirus	Central nervous defects, mental retardation, oculo-auditory lesions, hepatosplenomegaly, thrombocytopenia, chorioretinitis, pneumonitis, intrauterine growth retardation, hearing deficit
Varicella-Zoster	Microcephaly, cerebellar and cortical atrophy, ocular defects, cutaneous and musculoskeletal defects
Toxoplasmosis	Intrauterine growth retardation, icterus, hepatosplenomegaly, thrombocytopenia, mental retardation, hydrocephalus, microcephalus, chorioretinitis, cerebral calcifications
Venezuelan Equine Encephalitis	Spontaneous abortion, destruction of the cerebral cortex (hydroanencephaly), microphthalmia
Maternal diseases	
Diabetes ID	Increased incidence of congenital defects, congenital heart defects, caudal dysplasia or caudal regression syndrome
Epilepsy treated	Increased incidence of congenital defects, congenital heart defects, facial clefts, AED syndrome
Phenylketonuria	Spontaneous abortion, microcephaly, mental retardation, intrauterine growth retardation
Iodine deficiency	Growth retardation, mental retardation
Thyroid diseases with antibodies	Hypothyroidism
Virilizing tumors	Masculinization of female fetuses
Pheocromocitoma	Spontaneous abortion
Connettive's autoimmune diseases	Congenital heart block

When an association is suspected, the possibility that it is due to chance occurrence must always be considered. Tests of statistical significance and confidence intervals are used to this purpose. They indicate the probability that what we observed is only due to chance. When this probability is low (less than 5 or 1%), it is judged unlikely and therefore the observed association is considered real.

Confronting the problems

There are several methodological problems in conducting epidemiologic studies before firm conclusions can be reached. Only some of them commonly encountered in studies dealing with birth defects will be mentioned here.

Confounding. - An observed association may not be due to the exposure under study but to other factors associated both to the exposure and to the disease. These factors are named confounding factors. They are real risk factors of the disease under study and their distribution is different in exposed and unexposed subjects. Typical potential confounding factors in reproductive epidemiology are: maternal age, maternal disease status (i.e. epilepsy), life style (smoking, alcohol). For example, if smoking increases with the woman's age and we study Down's syndrome and smoking, a positive association will be found between smoking and Down's syndrome which is actually due to maternal age. When an association is suspected to be in fact due to confounding factors, particular approaches in study design or in data analysis can be taken to take it into account, given that information and knowledge on confounding factors are available.

Multiple comparison. - When several statistical tests are performed to test many different hypotheses in the same study, statistical significant associations can show up only for the effect of chance. This problem is more often encountered in studies where a large number of risk factors is investigated without a definite a priori hypothesis (fishing expeditions or exploratory studies). Conclusions from this type of studies must always be drawn with extreme caution and considered worth of further investigations.

Sample size and study power. - The conclusions of any type of study must be judged considering its sample size and power. Sometimes the absence of a statistical significant association between an exposure and a birth defect is due to the size of the study more than to the real absence of the association. If the sample size is too small, the power of the study to detect a difference becomes so tiny that only very large associations can be detected. On the other hand it is possible in theory to detect any difference, even trivial ones, between groups given a large enough sample size. This observation addresses the issue of the role of statistical analysis. Statistics allows the investigator to eliminate "chance" as a likely interpretation of the results of a study, but does not generally give insight into the mechanism of action. The conclusions of any study should in fact be based on biological, biochemical or other substantive grounds and be biologically plausible [16].

Methodological approaches have been developed (meta-analysis) [17] to cumulate studies carried out through comparable study designs. This allows to obtain larger sample sizes and more reliable estimates of the existence of the association. A meta-analysis of the many studies regarding Bendectin has been previously mentioned [10].

Criteria of causality

In order to establish a causal association between a teratogen and a birth defect, all or most of the following classic criteria of causality [18] should be met.

Specificity. - If the teratogen is associated only to one or a few specific birth defects, the possibility of a spurious association becomes smaller. Specificity of an association supports a causal interpretation but the lack of specificity does not negate it. The main weakness of specificity is that the concept is generally too simplistic. Multiple causes and effects are more often the rule than the exception.

For example the association between penicillamine and cutis laxa supports the teratogenicity of the drug. On the other hand there are well known teratogens, as high doses of ionizing radiation, which causes several different types of disease: mental retardation, leukemia, cataract and cancer of various districts [19].

Consistency among studies. - A second criterium that supports a causal interpretation of an association is the repeated observation of the association under different conditions of study. When this happens it is unlikely that methodological problems or systematic biases can influence the results of studies conducted in quite different contexts and perhaps different study designs.

Biological plausibility. - When a chemical or any other environmental factor caused a malformation in the experimental animals and/or the biological mechanism is understood or reasonably envisaged, the observation of an association in humans becomes more plausible. Biological plausibility is a time related concept in consideration of the evolution of scientific knowledge: indeed the observation of a seemingly implausible association may actually represent the beginning of an extension of our knowledge. However, the necessity that a biologically plausible association must exist before any association can be regarded as causal, indicates the

distinction between statistical significance and biological significance in epidemiologic studies. Although the statistical association must be present before any relationship can be said to exist, only biologically plausible associations can result in "biological significance" [11].

Strength of the association. - The larger the value of the relative risk, the less likely the association is to be spurious [20].

If the association between a teratogen is weak and the relative risk small (i.e. range 1.1-2), it is possible to think that the association is indeed due to unknown confounding factors and not to the teratogen under study. However weak associations may be due to misclassification of exposure or of disease; they may also indicate an overall low risk but the presence of a special subgroup at risk of disease within the exposed group.

Dose-response relationship. - If a factor is the cause of a disease, then the risk of developing the disease should be related to the degree of exposure to the factor: that is, a dose response relationship should exist. An observed dose response relationship makes a causal hypothesis more plausible. For instance the virilizing effect of androgens seems to be dose-dependent either in terms of disease severity or in terms of frequency of effects [21].

Gestation time of exposure

Teratogenesis resulting in a specific malformation may occur only during a rather limited period of development. The length of the "sensitive" period varies depending on the particular teratogenic process and the target organ of the teratogen.

The maximum sensitive period of some malformations is summarized in Table 2. If a possible teratogen is administered outside the period of organ susceptibility it cannot interfere with the process of development of one or more organs or systems unless it determines the regression of an already developed organ [4]. However, the incomplete knowledge on embryonic development suggests careful evaluations in considering the period of susceptibility to a specific teratogen [22].

Conclusions

A list of identified teratogens is summarized in Table 3. They are a limited number and, as already mentioned, explain only 6% of all birth defects. There is a substantial lack of knowledge in this area of reproductive epidemiology, worth of further reserch and resources considering the important social, emotional and public health impact of these adverse outcomes of pregnancy. Birth defects in developed countries are still an important cause of infant and child mortality which was only marginally influenced by the progress of postnatal treatment. Because of advances in molecular techniques, cytogenetic methods, statistical genetics and biochemical methods to measure exposure and genetic susceptibility, along with contributions from clinical medicine, embriology and other fields, progress is being made toward understanding the causes and mechanisms of birth defects in humans. The study of birth defects in the next years will be a major challenge for reproductive epidemiology.

Acknowledgements

Project supported by ASM-Associazione Italiana per lo Studio delle Malformazioni, by the FATMA project of the Consiglio Nazionale delle Ricerche-CNR - Contratto 9200179 PF41, and by the Centro Studi of the Ministero della Sanità.

Submitted on invitation.
Accepted on 25 September 1992.

REFERENCES

1. BRENT, R.L. 1986. Evaluating the alleged teratogenicity of environmental agents. *Clin. Perinat.* **24**: 609-613.

2. WILCOX, A.J., WEINBERG, C.R., O'CONNOR, J.F. *et al.* 1988. Incidence of early loss of pregnancy. *N. Engl. J. Med.* **319**: 189-194.

3. WARBURTON, D., STEIN, Z., KLINE, J. *et al.* 1980. Chromosome abnormalities in spontaneous abortions: data from the New York City study. In: *Human embryonic and fetal death.* I.H. Porter & E.B. Hook (Eds). Academic Press, New York.

4. INTERNATIONAL CLEARINGHOUSE FOR BIRTH DEFECTS MONITORING SYSTEMS. 1991. *Congenital malformations worldwide.* Elsevier Science Publishers BV, Amsterdam.

5. KALLÉN, B. 1988. *Reproductive epidemiology.* CRC Publisher, Boca Raton, Florida.

6. KALTER, H. & WARKANY, J. 1983. Congenital malformations. Etiologic factors and their role in prevention. *N. Engl. J. Med.* **308**: 424-431.

7. LENZ ,W. & KNAPP, K. 1962. Foetal malformations due to thalidomide. *Ger. Med. Mthly.* **7**: 253-258.

8. GOLDBERG, J.D. & GOLBUS, M.S. 1986. The value of case reports in human teratology. *Am. J. Obstet. Gynecol.* **154**: 479-482.

9. NEWMAN, S.A. 1990. Bendectin-birth defects controversy. *JAMA* **264**(5): 560.

10 HOLMES, L.B. 1983. Teratogen update: Bendectin. *Teratology* **27**: 277-281.

11. LILIENFELD, A.M. & LILIENFELD, D.E. 1980. *Foundations of epidemiology.* Oxford University Press, New York.

12. WERLER, M.M., POBER, B.R., NELSON, K. *et al.* 1989. Reporting accuracy among mothers of malformaed and nonmalformed infants. *Am. J. Epidemiol.* **129**: 415-421.

13. MACMAHON, B. & PUGH, T.F. 1970. *Epidemiology. Principles and methods*. Little Brown and Co., Boston.

14. MRC VITAMIN STUDY RESEARCH GROUP. 1991. Prevention of neural tube defects. Results of the Medical Research Council Vitamin Study. *Lancet* **338**: 131-137.

15. HOLTZMAN, N.A. & KHOURY, M.J. 1986. Monitoring for congenital malformations. *Annu. Rev. Public Health* **7**: 237-266.

16. SCHLESSELMAN, J.J. 1982. *Case Control Studies: design, conduct, analysis*. Oxford University Press, New York.

17. GLASS, G.V. 1976. Primary, secondary, and meta-analysis of research. *Educ. Res.* **5**: 3-8.

18. HILL, A.B. 1965. The environment and disease: association or causation? *Proc. R. Soc. Med.* **58**: 295-300.

19. LIONE, A. 1987. Ionizing radiation and human reproduction. *Reproduct. Toxicol.* **1**(1): 3-16.

20. CORNFIELD, J., HAENSZEL, W., HAMMOND, E.C. *et al.* 1959. Smoking and lung cancer: recent evidence and discussion of some questions. *J. Natl. Can. Inst.* **22**: 173-203.

21. PAGANO, M. & MASTROIACOVO, P. 1988. *La prescrizione dei farmaci in gravidanza*. McGraw-Hill, Milano.

22. IANNACCONE, P.M., BOSSERT, N. & CONNELLY, C.S. 1987. Disruption of embryonic and fetal development due to preimplantation chemical insults. A critical review. *Am. J. Obstet. Gynecol.* **157**: 476-484.

23. SMITH, D.W. 1982. *Recognizable patterns of human malformations*. Saunders W.B. Co., Philadelphia.