

4. No quantitative specifications have been presented to control when reprocessing will occur, that is; if the actives/inactives are too low/high. A maximum and minimum limit should be specified at which no reprocessing occurs, for both actives and inactives.
5. An in-process test should be developed to measure for completeness of solution/presence of microscopic particles in solution. USP test <641> would be appropriate.
6. The to-be-marketed formulation W1133-M-08-A contains a [redacted] overage of 4-hydroxyanisole. Overages for losses during manufacturing are acceptable provided a rationale is presented; however, overages are not allowed for losses during storage.
7. The photostability studies, which resulted in the label statement "store in carton", were performed under ambiguous conditions. The photostability studies should be performed according to ICH guidelines.

PHARMACOLOGY/TOXICOLOGY:

If a comparison of the impurity profiles between the new and previous suppliers of 4-hydroxyanisole reveals new unqualified impurities, it may be necessary to perform additional testing prior to approval.

CLINICAL:

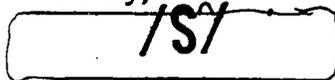
1. Safety concerning cutaneous adverse events needs to be demonstrated in population subsets other than Caucasians. This is important information for labeling.
2. Lack of data concerning [redacted] given the different mechanisms of pathogenesis of these lesions, would preclude its addition to the current indication.

Since the time available to complete the review of this application is limited, we will appreciate your prompt attention to the above issues. A final action on this application will follow the resolution of these issues, or may follow at any time if it becomes clear that all remaining critical issues can not be adequately resolved within this review cycle.

**APPEARS THIS WAY
ON ORIGINAL**

If you have any questions, contact Frank H. Cross, Jr., M.A., CDR, Senior Regulatory Management Officer, at (301) 827-2020. You are encouraged to discuss with us the issues identified above before investing resources in additional studies.

Sincerely,



12/23/1998

Robert DeLap, M.D., Ph.D.

Director

Office of Drug Evaluation V

Center for Drug Evaluation and Research

**APPEARS THIS WAY
ON ORIGINAL**



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

Food and Drug Administration
Rockville MD 20857

Date: December 1, 1998

DEC 1 1998

To: NDA 20-922 file

From: William C. Timmer, Ph.D.
Review Chemist, HFD-540

/S/

1 DEC 98

Through: Wilson H. DeCamp, Ph.D.
Chemistry Team Leader, HFD-540

/S/ 2/1/98

Subject: Addendum to Chemistry Review #1, dated 13-Nov-98

This addendum concerns the CMC issues presented on pg. 4 and again in the 'Draft Letter to the Applicant.' (The CMC issues are identical in both places in the review).

The CMC issues are to be re-grouped into two sections: NOT APPROVABLE and APPROVABLE.

NOT APPROVABLE:

No information regarding the manufacture of the 4-hydroxyanisole drug substance was available. Manufacturing information must be made available for review, whether it is included in the NDA itself or incorporated into a drug master file (DMF).

The container/closure DMFs were reviewed and found to be deficient. The DMFs are:

- DMF [redacted], held by the [redacted] for the cap, wiper, applicator rod and tip.
- DMF [redacted] held by [redacted] for the high-density polyethylene bottle.

APPROVABLE:

A range should be specified for the quantities of the individual excipients of the drug product.

No quantitative specifications have been presented to control when reprocessing will occur; that is, if the actives/inactive are too low/high. A maximum and minimum limit should be specified at which no reprocessing occurs, for both actives and inactives.

An in-process test should be developed to measure for completeness of solution/presence of microscopic particles in solution. USP test <641> would be appropriate.

The to-be-marketed formulation W1133-M-08-A contains a overage of 4-hydroxyanisole. Overages for losses during manufacturing are acceptable provided a rationale is presented; however, overages are not allowed for losses during storage.

The photostability studies, which resulted in the statement label "store in carton" were performed under ambiguous conditions. The photostability studies should be performed according to ICH guidelines.

cc: Orig. NDA 20-922
HFD-540/Division File
HFD-540/Division Director
HFD-540/PM/FCross *Director of Office*

HFD-540/MO/DCook
HFD-540/Pharm/ANostrandt
HFD-540/BioPharm/VTandon
HFD-805/Micro/PStinavage
HFD-540/Chemistry/WTimmer
HFD-540/ChemTmLdr/WHDeCamp

HFD-830/DD/CWChen
HFD-830/DDD/BDunn

APPEARS THIS WAY
ON ORIGINAL

Printed by Frank Cross, Jr.
Electronic Mail Message

Date: 10-Feb-1999 08:58am
From: Dan Boring
BORINGD
Dept: HFD-530 CRP2 S447
Tel No: 301-827-2396 FAX 301-827-2510

Subject: Re: NDA 20-922

Bill,

The proposed proprietary name SOLAGE' was found fully acceptable
by the Committee.

thanx,
dan

**APPEARS THIS WAY
ON ORIGINAL**

Printed by Frank Cross, Jr.
Electronic Mail Message

Sensitivity: COMPANY CONFIDENTIAL

Date: 05-Mar-1999 04:53pm
From: William Timmer
TIMMERW
Dept: HFD-540 CRP2 N224
Tel No: 301-827-2048 FAX 301-827-2075

TO: Frank Cross, Jr.
TO: Wilson DeCamp

(CROSSF)
(DECAMP)

Subject: Mequinol to 4-HA

Frank: On the label, for the package insert, and carton-container labeling, and medication guide, mequinol is 4-hydroxyanisole. Please replace, in the description section, **FROM:**

The chemical name for 4-HA, the hydroquinone, is 1-hydroxy

TO:

Mequinol is 4-HA, thehydroquinone or 1-hydroxy

OK? (Don't forget to remove the comma).

Accordingly, 4-hydroxyanisole should be replaced by mequinol throughtout the rest of the lableing, contingent upon approval from the other disciplines.

Bill

APPEARS THIS WAY
ON ORIGINAL



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

Food and Drug Administration
Rockville MD 20857

Division of Dermatologic and Dental Drug Products

Office of Drug Evaluation V
Center for Drug Evaluation and Research
Food and Drug Administration
9201 Corporate Boulevard, HFD-540
Rockville, MD 20850

FACSIMILE TRANSMISSION

DATE: March 30, 1999

TO: Donald Handley, Manager, Worldwide Regulatory Affairs
COMPANY: Bristol Myers Squibb
FAX #: 716-887-3638

MESSAGE: Please find attached to this facsimile transmission a copy of the Action Letter for NDA 20-922, Solagé (mequinol, 2%/tretinoin, 0.01%) Topical Solution.

Thank you.

FROM: Frank H. Cross, Jr., M.A., CDR
TITLE: Senior Regulatory Management Officer
PHONE #: 301-827-2063
FAX #: 301-827-2075/2091

THIS DOCUMENT IS INTENDED ONLY FOR THE USE OF THE PARTY TO WHOM IT IS ADDRESSED AND MAY CONTAIN INFORMATION THAT IS PRIVILEGED, CONFIDENTIAL, AND PROTECTED FROM DISCLOSURE UNDER APPLICABLE LAW. If you are not the addressee, or a person authorized to deliver the document to the addressee, you are hereby notified that any review, disclosure, dissemination, copying, or other action based on the content of this communication is not authorized. If you have received this document in error, please immediately notify us by telephone.



NDA 20-922

MAR 30 1999

Bristol-Myers Squibb Pharmaceutical Research Institute
Attention: Donald J. Handley, Manager, Worldwide Regulatory Affairs
100 Forest Avenue
Buffalo, New York 14213-1091

Dear Mr. Handley:

Please refer to your new drug application (NDA) dated December 30, 1997, received December 30, 1997, submitted under section 505(b) of the Federal Food, Drug, and Cosmetic Act for Solagé (mequinol, 2%/tretinoin, 0.01%) Topical Solution.

Please also refer to our Information Request Letter dated December 23, 1998.

We acknowledge receipt of your amendments and correspondence dated January 8 (two), February 5 (two), 10, 18, 22, March 2 (two), 4, 5, 9, 16, 22, 23, 25 (two) and 26, 1999. The user fee goal date for this application is March 30, 1999.

This new drug application provides for the use of Solagé (mequinol, 2%/tretinoin, 0.01%) Topical Solution for the treatment of solar lentigines.

We have completed the review of this application, as submitted with draft labeling, and it is approvable. Before this application may be approved, however, it will be necessary for you to submit the following:

Draft labeling for the drug product revised as recommended in the enclosed revised draft labeling (text for the package insert, text for the patient medication guide, immediate container and carton labels). Should additional information relating to the safety or effectiveness of this drug become available, revision of the labeling may be required.

We note that the product labeling issues that have not yet been resolved to our mutual satisfaction primarily concern reproductive toxicology and the product pregnancy category. We have discussed with you the potential value of additional preclinical research to help address some of these issues. Please contact us if we can provide further assistance, and submit for review any protocols you may develop for such additional studies.

We also acknowledge your Phase 4 commitments specified in your submission dated March 26, 1999. These commitments are described below:

1. Conduct an open label study with 200-300 subjects in ethnic groups of Asian, Latin/Hispanic, and of African descent with skin types II-V, to determine the safety of Solagé regarding pigmentary changes in the skin within 30 months of approval of NDA 20-922. Please submit the protocol to your IND for this product for our review three months prior to initiation of the study.
2. Conduct a 4-week bridging study in rabbits to demonstrate equivalency of the new mequinol to that used in previous studies within one year of approval of NDA 20-922. Alternatively, if analysis of reserve test article from long-term nonclinical studies reveals similar levels of the same impurity(ies) to that found in the new material, then those studies may be considered sufficient to qualify the impurity(ies). Please submit the protocol to your IND for this product for our review three months prior to initiation of the study.

Under 21 CFR 314.50(d)(5)(vi)(b), we request that you update your NDA by submitting all safety information you now have regarding your new drug. Please provide updated information as listed below. The update should cover all studies and uses of the drug including: (1) those involving indications not being sought in the present submission, (2) other dosage forms, and (3) other dose levels, etc.

Within 10 days after the date of this letter, you are required to amend the application, notify us of your intent to file an amendment, or follow one of your other options under 21 CFR 314.110. In the absence of any such action FDA may proceed to withdraw the application. Any amendment should respond to all the deficiencies listed.

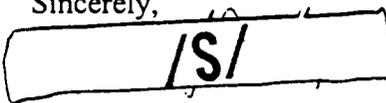
We will not process a partial reply as a major amendment nor will the review clock be reactivated until all deficiencies have been addressed.

Under 21 CFR 314.102(d) of the new drug regulations, you may request an informal or telephone conference with this Division to discuss what further steps need to be taken before the application may be approved.

The drug product may not be legally marketed until you have been notified in writing that the application is approved.

If you have any questions, contact Frank H. Cross, Jr., M.A., CDR, Senior Regulatory Management Officer, at (301) 827-2020.

Sincerely,



3/30/1997

Robert J. DeLap, M.D., Ph.D.
Director
Office of Drug Evaluation V
Center for Drug Evaluation and Research

Enclosure

**APPEARS THIS WAY
ON ORIGINAL**

16 Page(s) Redacted

Draft

Labeling

2 Page(s) Redacted

Draft

Labeling



Division of Dermatologic and Dental Drug Products
Office of Drug Evaluation V
Center for Drug Evaluation and Research
Food and Drug Administration
9201 Corporate Boulevard, HFD-540
Rockville, MD 20850

FACSIMILE TRANSMISSION

DATE: March 26, 1999 **Number of Pages (including cover sheet) - 1**

TO: Donald Handley, Manager, Worldwide Regulatory Affairs
COMPANY: Bristol Myers Squibb
FAX #: 716-887-3638

MESSAGE: Please commit to the following Phase 4 commitments for NDA 20-922, Solagé (mequinol, 2%/tretinoin, 0.01%) Topical Solution:

1. Conduct an open label study with 200-300 subjects in ethnic groups of Asian, Latin/Hispanic, and of African descent with skin types II-V, to determine the safety of Solagé regarding pigmentary changes in the skin within 30 months of approval of NDA 20-922. The protocol will be submitted to your IND for this product for our review three months prior to initiation of the study.

2. Conduct a 4-week bridging study in rabbits to demonstrate equivalency of the new mequinol to that used in previous studies within one year of approval of NDA 20-922. Alternatively, if analysis of reserve test article from long-term nonclinical studies reveals similar levels of the same impurity(ies) to that found in the new material, then those studies may be considered sufficient to qualify the impurity(ies). The protocol will be submitted to your IND for this product for our review three months prior to initiation of the study.

Thank you.

FROM: Frank H. Cross, Jr., M.A., CDR
TITLE: Senior Regulatory Management Officer
PHONE #: 301-827-2063
FAX #: 301-827-2075/2091

THIS DOCUMENT IS INTENDED ONLY FOR THE USE OF THE PARTY TO WHOM IT IS ADDRESSED AND MAY CONTAIN INFORMATION THAT IS PRIVILEGED, CONFIDENTIAL, AND PROTECTED FROM DISCLOSURE UNDER APPLICABLE LAW. If you are not the addressee, or a person authorized to deliver the document to the addressee, you are hereby notified that any review, disclosure, dissemination, copying, or other action based on the content of this communication is not authorized. If you have received this document in error, please immediately notify us by telephone.

20 Page(s) Redacted

Draft
Labeling

TITLE

Teratogenicity of N-(4-hydroxyphenyl)-all-trans-retinamide in rats and rabbits.

AUTHOR(S)

Kenel-MF; Kraye-JH; Merz-EA; Pritchard-JF

SOURCE (BIBLIOGRAPHIC CITATION)

Teratog-Carcinog-Mutagen. 1988; 8(1): 1-11.

INTERNATIONAL STANDARD SERIAL NUMBER

0270-3211

LANGUAGE OF ARTICLE

ENGLISH

ABSTRACT

N-(4-hydroxyphenyl)-all-trans-retinamide (HPR) has potential efficacy in the treatment of dermatologic, arthritic, and neoplastic disorders. The teratogenicity of such a compound is of special concern in light of the known adverse effects of retinoids, in general, on the developing conceptus. In these studies, Sprague-Dawley rats and New Zealand White rabbits were treated orally from gestation days 6 to 15 and 6 to 18, respectively, with 0, 20, 125, or 800 mg/kg/day of HPR. In rat fetuses, low incidences of hydrocephaly (mid- and high-dosage groups) were observed. Fetal tissue (ng/g) and maternal plasma (ng/ml) concentrations of HPR, its major metabolite (N-[4-methoxyphenyl] retinamide [MPR]) and retinol were determined in separate groups of similarly-treated rats 3 h following the last dose on gestation day 15. Fetal tissue concentrations of HPR and MPR were approximately one-half maternal plasma concentrations. A dose related reduction in maternal plasma and fetal tissue concentrations of retinol were also observed. In mid- and high-dosage rabbit fetuses, a dose-related increase in the incidence of dome-shaped head was observed. Subsequent skeletal evaluation revealed delays in skull bone ossification and a widening of the frontal and frontoparietal sutures. Microphthalmia was also observed in two high-dosage fetuses. A dose-dependent and statistically significant reduction in maternal plasma retinol levels was observed across all dosage groups.

MEDLINE ACCESSION NUMBER

88236321

APPEARS THIS WAY
ON ORIGINAL

Early Embryonic Cell Response in Retinoid-Induced Teratogenesis

D. M. Kochhar and H. Jiang

Thomas Jefferson University, Philadelphia, Pennsylvania

D. R. Soprano and D. C. Harnish

Temple University, Philadelphia, Pennsylvania

INTRODUCTION

The occurrence of a distinct pattern of dysmorphogenesis in human infants after prenatal exposure to therapeutic doses of 13-*cis* retinoic acid (Accutane, isotretinoin) has stimulated renewed efforts to search for mechanism and mode of action of this teratogen. Salient features of the retinoid syndrome include distinctive craniofacial and cardiovascular defects, absent or hypoplastic thymus, hydrocephalus, malformations of the derivatives of the rhombencephalic alar plate, and a small incidence of unusual limb defects (Table 1). Some of the surviving children, including those who have escaped disfigurement, appear to sustain intellectual deficits in later development (Adams, 1990).

All-*trans* retinoic acid (RA) and 13-*cis* retinoic acid are interconvertible isomers, both of which occur normally in the body as metabolites of vitamin A (retinol) and certain carotenoids. Endogenous RA has been suggested to participate in organogenetic pattern formation in normal vertebrate embryos through transcriptional control of important development genes such as the homeobox (*Hox*) genes, and this action is most likely mediated by its nuclear receptors (see review in Tabin, 1991). In utero exposure of early mouse embryos to a moderate dose of RA resulted in homeotic transformations in the axial skeleton and in a disorganization of

Table 1 Teratogenicity of Accutane (13-*cis* RA) in Human Embryos

System Defect
Craniofacial
Malformed external ears ✓
Midfacial and mandibular underdevelopment
Wide cleft palate ✓
Absent thymus
Cardiovascular
Various arch and septation anomalies
CNS
Hydrocephalus, derivatives of rhombencephalic alar plate malformed ✓
Limbs
Two cases of bone dysplasia

Note: 0.5–1.5 mg/kg daily for variable period, during the first 10 weeks after conception.

Source: Lammer et al. (1985); Lammer and Armstrong (1991).

the hindbrain segmentation, both of which were accompanied by ectopic or heterotopic expression of certain *Hox* genes (Kessel and Gruss, 1991; Morriss-Kay et al., 1991). Introduction of exogenous RA into the anterior margin of the chick embryo wing bud also ectopically activated members of the *Hox-4* gene cluster, and this occurred in a precise spatiotemporal pattern which correlated with the subsequent mirror-image duplication of the digits (Eichele, 1989; Izpisua-Belmonte et al., 1991; Nohno et al., 1991).

The notion that during embryonic development a disruption in the homeobox gene function culminates in a distinct pattern of malformations was strengthened by a recent report (Chisaka and Capecchi, 1991). Newborn homozygous transgenic mice in which the *hox-1.5* gene was ablated through gene targeting presented a number of abnormalities in craniofacial and heart development and died shortly after birth (Table 2). The tissues derived from all pharyngeal arches, including the thymus, were defective; collectively, the pattern of defects reminded the authors of equivalence with the features of the DiGeorge syndrome in humans. It is also interesting to note that there are many similarities between components of the DiGeorge syndrome and those of retinoic acid embryopathy. Such similarities, however, may only be superficial and require cautious interpretation. It has been noted that the DiGeorge syndrome is autosomal dominant, whereas the mouse *hox-1.5* phenotype is recessive, and that the human syndrome is associated with deletions and translocations of chromosome 22

Table 2 *Hox-1.5* Gene Ablation in the Mouse (Chromosome 6)

Congenital Defects	Normal expression
Craniofacial defects, multiple defects in derivatives of the pharyngeal arches, heart, and great vessels, patterning of tissues in the neck, throat, and tongue. Neonatal death. Autosomal recessive.	Hindbrain and spinal cord, spinal ganglia, pharynx, aortic trunk, thyroid, lungs, stomach, kidneys
Likely equivalent to DiGeorge's syndrome	
Reservations: Autosomal dominant. Chromosome 22 deletions and translocations. Human homologue to <i>hox-1.5</i> maps to Chromosome 7)	

Source: Chisaka and Capecchi (1991).

rather than the chromosome 7 where the homolog to *hox-1.5* is mapped (Chisaka and Capecchi, 1991). Also, the pattern of expression of *hox-1.5* gene in normal mouse embryos is much more widespread in tissues and organs, and therefore incongruous with the more restricted pattern of defects found after its ablation (Gaunt, 1987). Apparently, there are other molecules and other factors in the embryo that influence or determine whether or not an initial disruption in the function of one of the homeobox genes will culminate in a teratogenic outcome.

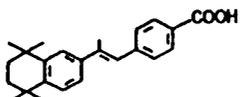
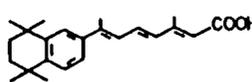
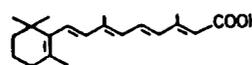
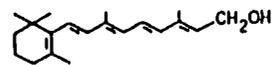
In this chapter, we summarize our attempts to study the process of teratogenesis in mice upon exposure to an oral RA dose. First, we have compared the teratogenic potency of RA with a few synthetic retinoids in a standardized teratology protocol to discern the dependence of teratogenic activity on the chemical structure of the retinoid molecule and to determine the extent to which pharmacokinetic parameters are responsible for the high potency of certain synthetic compounds in the mammalian embryo. Using the same teratology protocol, we further investigated the minimum duration of in utero exposure to RA required for the irreversible defective development of the target organ; this information would delimit the postexposure period during which to seek the earliest susceptible events that precede teratogenesis. Second, we describe our recent results on the ability of transplacentally derived RA to rapidly induce transcription of one of its own receptors (RAR- β) in the target organs, specifically in the limb bud and the craniofacial region, raising the likelihood that this receptor isoform is a mediator of the onset of teratogenesis. Finally, based on the assumption that there may be a common denominator among a cascade of receptor-mediated events in multiple target organs of the retinoid-exposed embryo, we present preliminary results on changes induced in the activity of tissue transglutaminase in the potentially abnormal limb buds during the early susceptible phases of development.

TERATOGENIC POTENCY OF RETINOIDS

The pattern of defects in animal embryos is dependent on the developmental stage at which RA is administered; a fact not yet demonstrable in the Accutane-exposed human fetuses (Shenefelt, 1972; Lammer et al., 1985). The teratology protocol we have used comprises morphological evaluation of near-term fetuses of mice exposed to a single oral dose of the retinoid on the morning of day 11 of gestation; incidence of resorption, growth retardation (fetal weight), palatal cleft, and limb reduction deformity are consistent and reliable endpoints with which to compare relative teratogenic potencies of these compounds (Kochhar, 1973; Kochhar et al., 1984). In this assay, RA was about four-fold more active than retinol, whereas tetrahydrotetramethylnaphthalenyl-propenyl benzoic acid (TTNPB) was almost 700-fold more active than RA (Table 3).

In resolving mechanisms of morphogenetic activity of retinoic acid, it would be instructive to know the basis for differential teratogenic activity of the retinoid analogs. Two explanations have been proposed. One, that the potency of an analog may depend on its binding affinity to nuclear receptors, RARs and RXRs. In addition to the nuclear receptors, there are two distinct cytoplasmic proteins, cellular retinoic acid-binding proteins

Table 3 Teratogenic Potencies of Synthetic Retinoids Relative to RA in Pregnant ICR Mice

RETINOID	STRUCTURE	TERATOGENIC POTENCY
Ro 13-7410 (TTNPB)		700 X
Ro 13-6307		40 X
RA (Retinoic Acid, all-trans)		1 X
Vitamin A (Retinol)		0.25 X

Source: Kochhar (1973), Kochhar (1987), Kochhar et al. (1988), Kochhar and Penner (1992).

(CRABP) I and II, that specifically bind to RA, but their functions are not entirely clear (Balling, 1991). In this regard, recent investigations have shown that although an active analog must possess an acidic end group in its structure to bind to the nuclear and cytoplasmic components, there appeared to be no direct correlation between binding and teratogenicity (Sani and Hill, 1976; Crettaz et al., 1990; Howard et al., 1990; Willhite, 1990). For example, the affinity of TTNPB for RAR- α , RAR- β , or CRABP was only 50–70% of that for RA (Crettaz et al., 1990; Howard et al., 1990). These findings suggest that other factors must account for the relative higher potency of the arotinoid. The arotinoids are known to be less labile to metabolic degradation than are the natural retinoids (Loeliger et al., 1980). It is also possible that their ability to activate receptor function may be, for some unknown reason, greater than that of RA.

A second explanation for differences in potency of retinoids deals with factors that determine the access to the mammalian embryo of pharmacologic agents delivered to the mother. It is now well established that administration of RA to pregnant animals results in its rapid transfer across the placenta to the developing embryo (Kochhar, 1976; Kochhar et al., 1987; Satre and Kochhar, 1989). The fact that 13-*cis* RA was less teratogenic than all-*trans* RA in rodents was shown to be associated with pharmacokinetic differences between isomers. Transplacental passage of 13-*cis* RA in the mouse was less efficient than all-*trans* RA. This resulted in exposure of the embryo to effectively lower concentrations of an otherwise active teratogen (Crech-Kraft et al., 1987). On the other hand, the high potency of arotinoids such as TTNPB remains to be explained.

In a recent study, we investigated how the transplacental transfer and embryonic concentrations of one of the potent synthetic retinoids compared with that of RA. We employed Ro 13-6307, an aromatic retinoid with a triene side chain, since it has been shown to possess lower toxicity than TTNPB (Loeliger et al., 1980; Howard et al., 1987). Reduced toxicity increased the likelihood of obtaining a sufficient number of viable embryos for quantitative analysis. Using the same teratology protocol as above, we found that this retinoid was 40-fold as active as RA (see Table 3) (Kochhar and Penner 1992). In view of the fact that Ro 13-6307 was much more teratogenic than RA, we inquired if pharmacokinetic parameters such as time- and dose-dependent concentrations reaching the embryo could account for this increased activity. We employed HPLC to identify and quantify the analog in the embryo for a period of 4 h after oral administration of teratogenic doses. We found that after a low oral dose of 1 mg/kg Ro13-6307, embryonic concentrations did not rise above 5 ng/g—our limit of detection, yet this treatment severely malformed more than 60% of the exposed fetuses. This finding sets Ro 13-6307 quite apart from conventional retinoids such as RA, retinol, and acitretin, which are known to

precipitate teratogenesis only if their concentrations in the embryo exceeded several hundred nanograms per gram value (see chap. 40). Since the morphogenetic activity of Ro 13-6307 is disproportionately in excess of its levels in the mouse embryo, obligatory mediation by the receptors and/or binding proteins is likely involved. It will be interesting in future studies to analyze the binding activity of Ro 13-6307 to the nuclear receptors and to CRABPs.

EARLY EMBRYONIC RESPONSE

In searching for causative mechanisms of teratogenesis, detailed studies on fetal phenotype are useful but less informative than those analyzing embryos shortly after RA exposure. The earliest changes that have been described were in cell cycle parameters of susceptible cell populations such as in the neural tube, facial processes, and the limb bud (Langman and Welch, 1967; Kochhar, 1968, 1977); this effect triggered other events, including unscheduled cell death, which disrupted the program of cell differentiation and organogenesis. Morphological evidence has also implicated aberrant development of the cranial neural crest cells, and derangements in cell-cell and cell-substratum interactions (Morriss, 1973; Kwasigroch and Kochhar, 1975; Morriss and Thorogood, 1978; Webster et al., 1986; Yasuda et al., 1989; Alles and Sulik, 1990). Since some of these early microscopic changes may be reversible owing to regulatory mechanisms that operate in early embryogenesis, the true significance of some of the altered events in teratogenesis is not fully established.

As a prerequisite to distinguishing primary causative mechanisms, we attempted to define the minimum duration of RA exposure necessary for irreversible determinative events to occur. We chose limb bud as the target organ, since our standard teratology protocol results in almost 100% incidence of limb reduction (phocomelia) anomalies (Kochhar, 1973, 1977). Pregnant DBA/2J mice were killed at hourly intervals after an oral RA dose of 100 mg/kg on day 11 of gestation, and embryos were removed and their limb buds explanted in organ culture where their growth and development were monitored (Kochhar and Aydelotte, 1974). Using parameters of growth, aberrant morphogenetic pattern, and frequency of occurrence, an index of limb deformity was obtained which was plotted as a function of the duration of the in utero RA exposure (Fig. 1) (Kochhar 1982; and unpublished). It became evident that the limb bud sustained only a minor morphogenetic effect during the first 2 h of exposure; the effect was progressively more severe with further exposure, so that the limb deformity index was 50% at 6 h. A 10- to 12-h exposure resulted in full phenotypic effect (Index = 100%), which could be equated with phocomelic appearance of the limbs at birth (Fig. 1). Hence, we concluded that developmen-

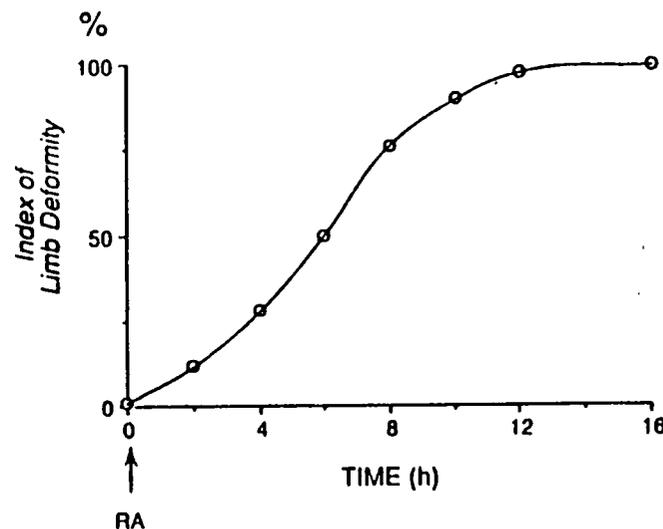


Figure 1 The index of limb deformity (growth, morphogenetic pattern, frequency of occurrence) is plotted as a function of time elapsed between the administration of RA to pregnant DBA/2J mice (100 mg/kg RA, day 11 of gestation) and the removal of embryos from the mother. The limb buds were explanted in organ culture to monitor their developmental potential. The developmental events susceptible to change by RA were initiated within 2 h of exposure and required about 10–12 h of exposure to produce the abnormal phenotype observed at birth. (From Kochhar and Aydelotte [1974] and Kochhar [1982].)

tal events susceptible to change by RA were initiated within 2 h of exposure and required no more than 12 h of exposure to produce the defect observed at birth.

This time frame of teratogenic action is coincident with the kinetics of appearance and attainment of peak RA concentrations within the limb bud following the teratogenic dose of RA (Fig. 2) (Satre and Kochhar, 1989). It also follows closely the profile of RA generated in the embryo after administration of a teratogenic dose of retinol, a fact that supports the contention that metabolic conversion of retinol to RA is obligatory in inducing the teratogenic response (Fig. 3) (Kochhar et al., 1988).

The discovery of retinoic acid receptors and the finding of endogenous RA in measurable quantities in the developing tissues have great promise in providing new insights into teratogenic mechanisms. All three genes that encode the nuclear receptors (RAR- α , - β , and - γ) are actively transcribed in the embryo and their mRNAs show distinct spatial and temporal distribution during development (see Tabin, 1991, for review). Although no spe-

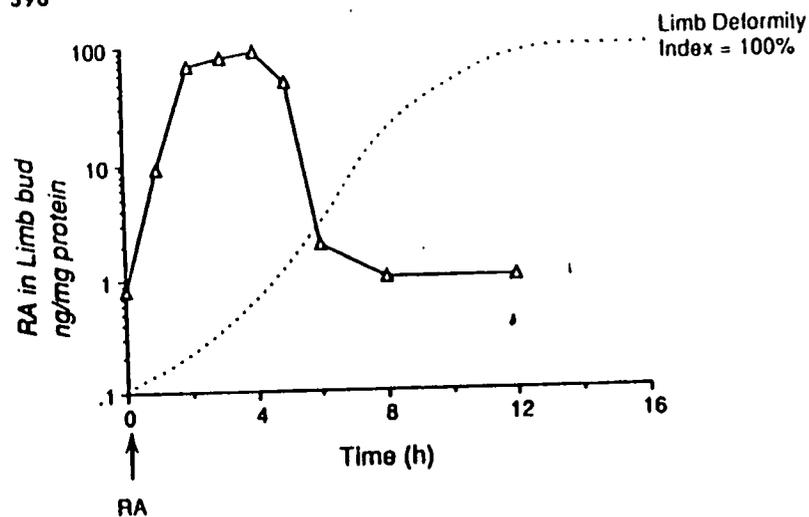


Figure 2. Concentrations of RA in the limb buds (ng/mg protein) of embryos from ICR mice are plotted as a function of time after an oral dose of 100 mg/kg RA on day 11 of gestation. The peak concentration occurred by 2 h and declined to basal level by 6 h of exposure. (From Satre and Kochhar [1989].)

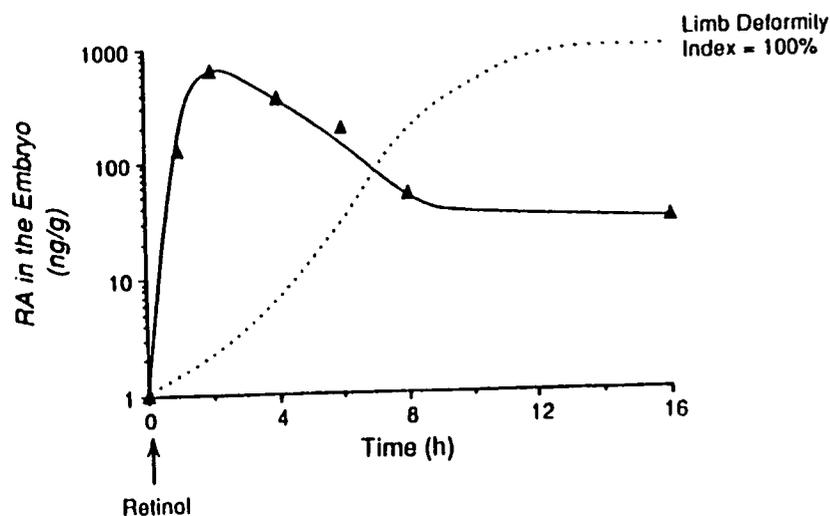


Figure 3. Concentrations of RA in the whole embryo (ng/g wet weight) are plotted as a function of time after an oral dose of retinol (200 mg/kg) administered to ICR mice on day 11 of gestation. (From Kochhar et al. [1988].)

cific insights into the role of any of the receptors in teratogenesis have emerged from the *in situ* localization studies, it has been reported that a teratogenic dose of RA resulted in a transient increase in RAR- β mRNA in the embryo without inducing any change in the basal level of RAR- α mRNA (Harnish et al., 1990). Since RAR- β has been suggested to have a role in morphogenetic cell death, and is now known to exist in three major isoforms, we inquired in an ongoing study if any of the RAR- β isoforms were preferentially activated in potentially abnormal limb buds during the determinative period of teratogenesis.

Total RNA extracted from control and RA-treated limb buds was hybridized with isoform-specific cDNAs in an RNase protection assay, and the signal was quantitated by densitometer (Harnish et al., unpublished data). While no effect of RA exposure was detected on the levels of $\beta_{1,3}$ mRNA, β_2 mRNA showed a rapid increase in the limb bud by 3 h of exposure to a level that was 16-fold greater than the control (Fig. 4). This increase was

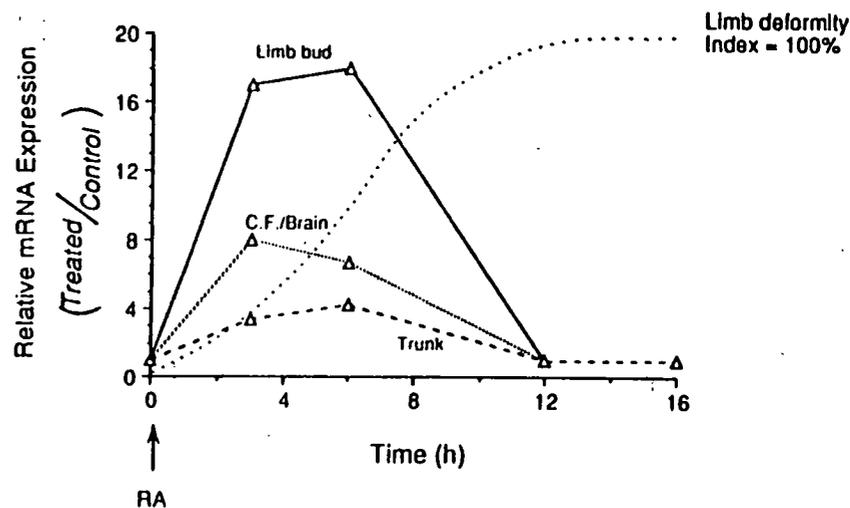


Figure 4. RAR- β_2 mRNA levels in the limb bud, craniofacial organs (CF) and brain, and the remaining body (Trunk) are plotted as a function of time after an oral dose of RA (100 mg/kg on day 11 of gestation). The levels of RAR- β_2 mRNA were determined by RNase-protection assay and quantitated by densitometry. The relative level of transcripts in each tissue was calculated with the control level at time 0 h set arbitrarily at 1. The treated limb buds show an induction response which is about 16-fold higher than the untreated limb buds at 3 and 6 h after dose. The β_2 mRNA levels return to basal level by 12 h after the dose. (From Harnish et al., unpublished.)

maintained until 6 h before declining to the low, steady-state level by 12 h of exposure. The increase was also detected in the embryonic brain and craniofacial complex, which amounted to about eight-fold induction over the control. Only a minor increase was found in the rest of the embryonic body (Fig. 4). This transient increase in β_2 mRNA in the susceptible tissues occurs at the time before a major increase in the limb deformity index is observed and coincides with the time of onset of detectable levels of cell death in the affected tissues, suggesting a critical role for this receptor isoform in teratogenesis. We have now also found a similar magnitude of β_2 mRNA induction in the target organs of embryos exposed to a teratogenic dose of retinol. Since the receptor is RA specific, we believe that this induction must have occurred through metabolic conversion of retinol to RA.

In certain cell types, investigators have observed a specific and early effect of retinoic acid on transcriptional activation of the gene for the peptide cross-linking enzyme transglutaminase (Uhl and Schindler, 1987; Chiocca et al., 1988). Tissue transglutaminase is known to be active in cells that undergo morphogenetic cell death, where it apparently cross-links membrane proteins, and thus prevents nuclear debris from leaking out and disrupting normal developmental program. Induction and activation of tissue transglutaminase (TGase) has been previously observed in hepatocytes undergoing programmed cell death (apoptosis) and in retinoic acid-induced differentiation of peritoneal macrophages and several myeloid cell lines in culture (Fesus et al., 1987; Chiocca et al., 1988; Piacentini et al., 1991). Since RA as a teratogen perturbs cell differentiation and induces cell death in affected embryos, this study was undertaken to determine if it also altered TGase activity in embryonic target tissues (Jiang and Kochhar, 1992). In a preliminary study, pregnant ICR mice were given an oral dose of RA (100 mg/kg) or vehicle alone on day 11 of gestation (plug day = day 0); maternal liver and embryonic limb buds were obtained at 1.5, 3.0 and 12.0 h after treatment. The TGase activity was assayed in tissue homogenates as the Ca^{2+} -dependent incorporation of [^3H]putrescine into dimethyl casein. The TGase activity in the maternal liver was very high (>900 pmol/min/mg protein) and remained unaffected by RA. The TGase activity in the limb buds (66.4 ± 21.5 pmol/min/mg protein) experienced a three-fold increase to 210.7 ± 59.0 pmol/min/mg ($p < 0.05$) by 3 h after RA treatment (Fig. 5). The enzyme activity returned to the level of control limb buds by 12 h after treatment (72.0 ± 19.2 pmol/min/mg; $p > 0.5$). The peak TGase activity corresponds in time to the peak accumulation of RA in the limb bud RA exposure (2–3 h) as well as to the first onset of cell death in the limb bud core mesenchyme (2–4 h). The induction of TGase activity in the limb bud was found to be dose-dependent ($r = 0.879$, $0 < 0.005$); upon exposure to the lower RA

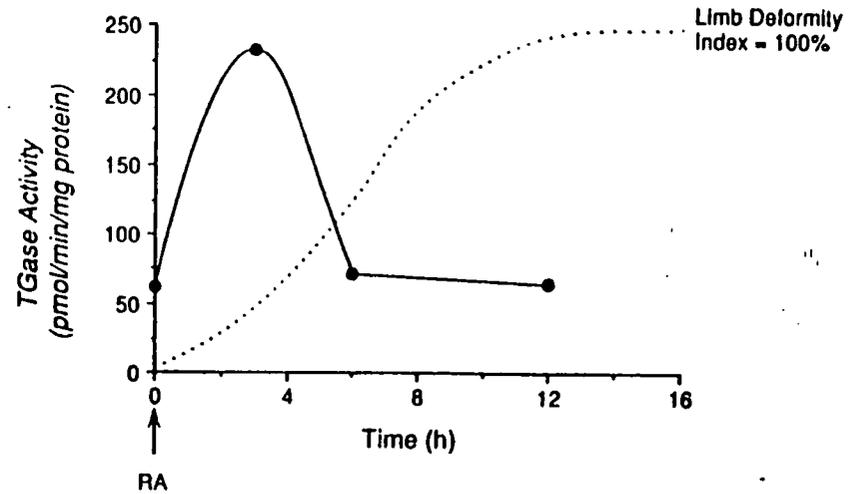


Figure 5 Tissue TGase activity (pmol/min/mg protein) in limb buds of embryos is plotted as a function of time after an oral dose of 100 mg/kg RA on day 11 of gestation. The enzyme activity increases 3-fold over the basal level by 3 h of exposure and returns to the basal level by 6 h of exposure. (From Jiang and Kochhar, 1992).

dose of 50 mg/kg, the activity increased at 3 h after dosing to an intermediate level of 134.2 ± 24.9 pmol/min/mg ($p < 0.05$). It is suggested that the induced TGase may be directly involved with some of the dysmorphogenic events in the RA-treated embryos. Further studies are needed to clarify the role of TGase and identify its substrate(s) in the limb bud.

ACKNOWLEDGMENTS

Original observations included in this chapter were supported by NIH grants HD-20925 (D.M.K.), DK-41089 and HD-27556 (D.R.S.), and a March of Dimes award 15-0059 (D.M.K.).

Assistance by John Penner for the illustrations and by Wanda Evans for word processing is gratefully acknowledged.

REFERENCES

- Adams, J. High incidence of intellectual deficits in 5-year old children exposed to isotretinoin *in utero*. *Teratology*, 41:614 (1990).
- Alles, A. J., and Sulik, K. K. Retinoic acid-induced spina bifida: Evidence for a pathogenetic mechanism. *Development*, 108:73-81 (1990).

Establishing the Timing of Implantation in the Harlan Porcellus Dutch and New Zealand White Rabbit and the Han Wistar Rat

DEBBIE A. GARSIDE,¹ ALLEN CHARLTON, AND KATE J. HEATH

Medicines Safety Evaluation Division, Glaxo-Wellcome Research and Development, Park Road,
Ware, Hertfordshire SG12 0DP, United Kingdom

Received November 20, 1995

The purpose of this study was to establish the timing of the onset of implantation in both the Harlan Porcellus Dutch and New Zealand White rabbit and the Han Wistar rat. Implantation was initiated on Day 5 (rat) and 7 (rabbit) and established by Day 7 and 8 of gestation in the rat and rabbit, respectively. Recent guidelines on toxicity testing during embryo-fetal development studies require that maternal exposure to pharmaceutical compounds does not occur until after implantation has taken place. In order to ensure that this is the case, female Harlan Porcellus Dutch and New Zealand White rabbits and Han Wistar rats were sacrificed on different days of gestation, over the expected periods of implantation. The presence of preimplantation blastocysts in the uterus was investigated, and evidence of established implantation sites was assessed. © 1996 Academic Press, Inc.

INTRODUCTION

The International Committee of Harmonisation has issued current guidelines on the detection of toxicity to reproduction (Commission of the European Committee for Proprietary Medicinal Products, 1993). These require that when studying the effects of medicinal products on embryo-fetal development, maternal exposure is initiated after implantation. It was therefore considered pertinent to establish the timing of implantation in the strains of rabbit and rat routinely employed in these studies.

Implantation in the rabbit is reported to occur around Day 7 of gestation (Lutwak-Mann, 1962) and in the rat to be late Day 5 (Hebel and Stanberg, 1986). Hence the period investigated in these studies was between Days 6 and 8 of gestation for the rabbit and Days 5 and 7 in the rat. The decidual reaction of implantation is normally induced by the implanting blastocyst (Krehbiel, 1937) and is characterized by increased uter-

ine vascular permeability and stromal edema. This increased permeability can be detected using the vascular leakage of the macromolecular dye, Pontamine blue (Psychoyos, 1971). The Pontamine blue reaction (PBR), therefore, results in the blue staining of all implantation sites present in the uterus within a day of implantation initiation.

This method has been widely used in the rat (Lundkvist and Ljungkvist, 1977), but there appears to be no published data on its use in the rabbit. Additionally, therefore, we investigated the use of Pontamine blue in the detection of implantation sites in the rabbit.

MATERIALS AND METHODS

Rabbit

Nine Harlan Porcellus (Firgrove Farm, Sussex) Dutch and nine Harlan Porcellus New Zealand White rabbits of at least 3.5 months old were housed singly. The females were fed approximately 120 g/animal/day of rabbit diet (Special Diets Services Ltd.) and water from the normal domestic supply *ad libitum*. Room temperature was maintained at 18°C and 45-70% relative humidity with a 12-hr light:dark cycle.

The females were mated with stock males and on confirmation of a positive mating (Day 1 of gestation), randomly allocated to three groups as follows:

Group	Number of Dutch females	Number of NZW females	Gestational day of sacrifice
1	3	3	6
2	3	3	7
3	3	3	8

Each female in each group was given 2.0 ml, 1.0% Pontamine blue in 0.9% sodium chloride, *iv*, 60 min prior to the time of sacrifice. The timing of sacrifice coincided with the normal dosing time for routine studies. As indicated above, the groups were sacrificed on

¹To whom correspondence and reprint requests should be addressed.

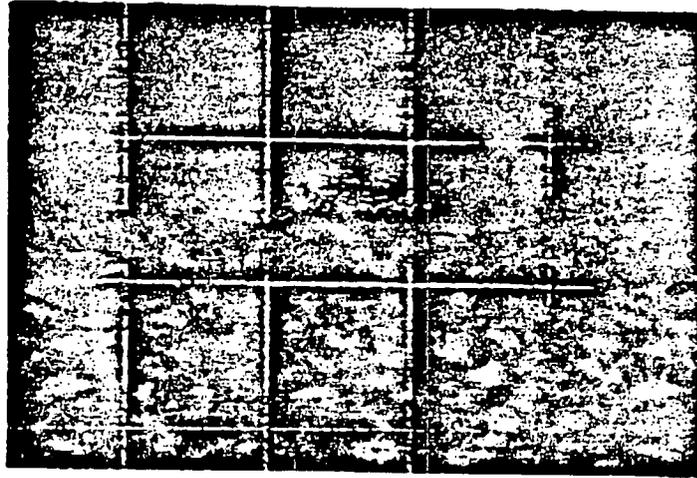


FIG. 1. Unattached embryos (arrowheads) present in the rabbit uterine lumen on Day 7 of gestation ($\times 100$). These were all at the blastocyst stage of development, and located in both strains of rabbit.

different days of gestation, thereby covering the estimated period of implantation in the rabbit.

All females were sacrificed by an iv injection of sodium pentobarbitone and as soon after death as possible, the reproductive tract was removed. The internal luminal surface of both uterine horns was exposed by cutting longitudinally along the antimesometrial line, and the number of implantation sites (stained blue) was counted. The number of corpora lutea present on the corresponding ovaries was also noted.

Rat

Fifteen female Han Wistar rats (Glaxo, Herts) of at least 8 weeks old were caged in pairs, on a 12:12 hr light:dark cycle with a temperature of 21°C and relative humidity of 45–70%. Rat and mouse Expanded No. 3 Diet (Special Diets Services Ltd) and water from the normal domestic supply were provided *ad libitum*. The females were mated with stock males and on confirmation of a positive mating (Day 1 of gestation) randomly allocated to three groups as follows:

Group	Number of Han females	Gestational day of sacrifice
1	5	5
2	5	6
3	5	7

All females were killed by inhalation of carbon dioxide, at a time in the morning coinciding with the normal dosing time for routine studies. However, to confirm whether implantation is initiated late on Day 5 of gestation, two females sacrificed on this day were done so

at 1600 hr. Additionally, this procedure was repeated for those sacrificed on Day 6. As soon after death as possible, the reproductive tract was removed and processed as described for the rabbit. For those females sacrificed on Day 5, the uteri were flushed with physiological saline prior to processing to ensure collection of unattached embryos.

RESULTS

Rabbit

The presence of embryos and implantation sites on different days of gestation, for the two strains, is shown in Table 1. Corpora lutea were present in all females, except one in Group 3, which was presumed not to be pregnant.

Group 1: Females sacrificed on Day 6 of gestation. In all females, no changes were detected to the uterine lumen by Pontamine blue. Additionally, the Fallopian tubes were flushed with physiological saline and em-

TABLE 1
Presence of Embryos (Blastocysts) and Implantation Sites in the Rabbit on Days 6–8 of Gestation

Uterine observation	Dutch rabbit: day of gestation			NZW rabbit: day of gestation		
	6	7	8	6	7	8
Embryos present	No	Yes	Yes	No	Yes	Yes
Implantation sites present	No	No	Yes	No	No	Yes

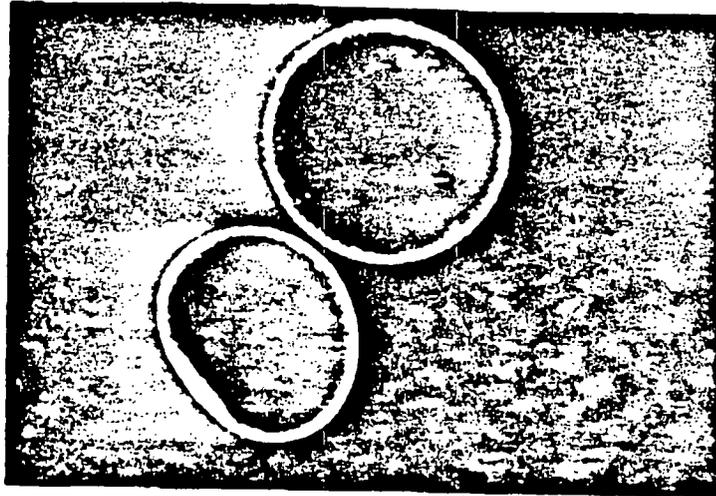


FIG. 2. A rabbit blastocyst recovered from the uterus on Day 7 of gestation, showing the embryonic disc (OE) ($\times 100$).

bryos at the morula stage of development were recovered.

Group 2: Females sacrificed on Day 7 of gestation. No implantation sites were observed by Pontamine blue staining. Embryos at the blastocyst stage of development were located in the uterine lumen, but these were unattached to the uterine epithelium (Figs. 1 and 2).

Group 3: Females sacrificed on Day 8 of gestation. Implantation sites were apparent by Pontamine blue staining (Fig. 3) in all females except one, which was

presumed not to be pregnant as corpora lutea were not present. All sites were established implantations, each supporting embryos developing normally. The average implantation site diameter was 0.5 cm.

Rat

The presence of embryos and implantation sites on different days of gestation, is shown in Table 2. Corpora lutea were present in all females.

Group 1: Females killed on Day 6 of gestation. In all females, implantation sites were not present and no

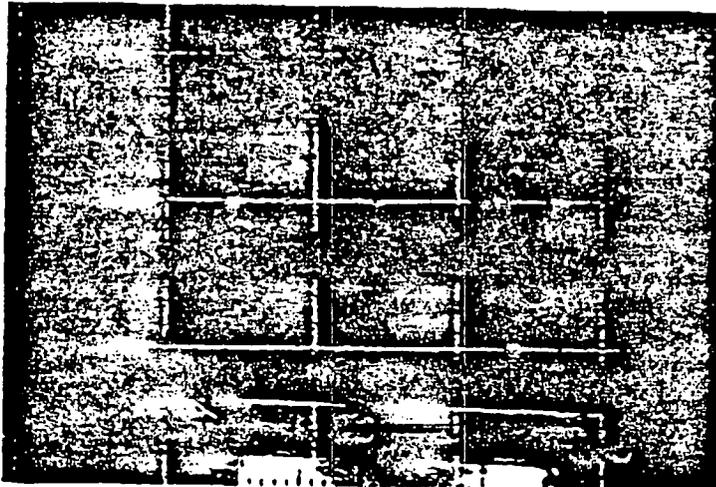


FIG. 3. Implantation sites (arrowheads), stained blue with Pontamine blue, present in the uterus of both strains of rabbit on Day 8 of gestation.

TABLE 2
Presence of Embryos (Blastocysts) and Implantation Sites in the Rat on Days 6-7 of Gestation

Uterine observation	Day of gestation		
	5	6	7
Embryos present	Yes	Yes	Yes
Implantation sites present	No	Yes	Yes

changes were located in the uterine lumens examined. On flushing, embryos were recovered at the blastocyst stage of development, which is the embryo stage necessary to coincide with implantation. Blastocysts recovered from rats killed on the morning of Day 5 possessed an intact zona pellucida (outer shell) (Fig. 4). Blastocysts recovered from rats killed in the afternoon were expanded and showed no evidence of a zona pellucida (Fig. 5). This is consistent with blastocyst physiology at the time of attachment to the endometrium (Cole, 1967; Kaufman, 1992).

Group 2: Females killed on Day 6 of gestation. Implantation sites were present in all females. These were obvious macroscopically in all females killed in the afternoon, but were only visible in those killed in the morning when the uterine lumen was exposed (Fig. 6). At the site of all implantations, there was a concentration of endometrial blood vessel network, indicating the initial stages of implantation.

Group 3: Females killed on Day 7 of gestation. In all females implantation sites were observed, which were established sites of implantation, supporting



FIG. 4. Unattached rat blastocyst present in the uterine lumen on the morning of Day 5 of gestation ($\times 200$). The zona pellucida (1) is intact.

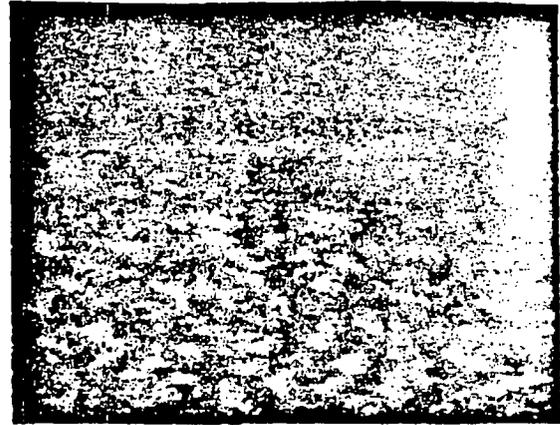


FIG. 5. Unattached rat blastocyst present in the uterine lumen late afternoon of Day 5 ($\times 200$). The zona pellucida is no longer present.

early embryos possessing developing amniotic cavities (Fig. 7).

DISCUSSION AND CONCLUSIONS

Implantation is initiated on Day 7 of gestation and established by Day 8 in both the *Harian Porcellus NZW* and Dutch rabbits. This is comparable with previous reports (Enders, 1976; Wimsatt, 1975) which indicate that implantation in the rabbit begins on Day 7 and is completed by Day 10 of gestation.

In the rat, implantation is initiated late on Day 5 and is completed by Day 7 of gestation. This is comparable to previous reports (Pincus, 1936; Nicholas, 1962).

Additionally, it is reported that implantation in the rabbit consists of two phases. The first occurs on Day 7-8 and is characterized by blastocyst adhesion to the



FIG. 6. Implantation sites (1) present in the rat uterus on the afternoon of Day 6 of gestation.

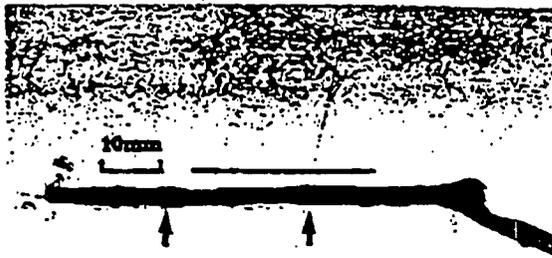


FIG. 7. Implantation sites (!) present in the rat uterus on Day 7 of gestation.

uterus and actual penetration of the uterine capillaries. The second phase covers the period from Days 8.5 to 10, during which time the formation of a definitive chorionicallantoic placenta is established.

The data presented here indicate that PBR is able to identify these two phases of implantation in the rabbit, as it does in the rat. This is evident by the fact that there was no staining present early on Day 7, when only blastocyst alignment in the uterus had occurred, but PBR was present on Day 8. At this time, blastocyst penetration of the capillaries had taken place, allowing vascular leakage of the dye and therefore a blue staining.

Implantation in the rabbit is obviously a complex series of events, taking place over a period of 2 days, which can be identified successfully with Pontamine blue. However, it must be noted that normal implantation in both the rat and the rabbit can exhibit variation, within as well as between individual females.

In conclusion, providing that dosing is started on or after Day 8 of gestation in both the Harlan Porcellus NZW and Dutch rabbits, and on or after Day 7 in the

Han Wistar rat, maternal/embryonic exposure to medicinal products will not occur until after implantation.

ACKNOWLEDGMENT

The authors thank the technical staff of the department for help with animal care and procedures.

REFERENCES

- Cole, R. J. (1967). Cincmicrographic observations on the trophoblast and zona pellucida of the mouse blastocyst. *J. Embryol. Exp. Morphol.* 17, 481-490.
- Commission of the European Communities for Proprietary Medicinal Products (1993). Ref.III/3387/93. *Detection of Toxicity to Reproduction for Medicinal Products* (Guideline prepared within the International Conference on Harmonisation (ICH) process).
- Endera, C. E. (1976). Anatomical aspects of implantation. *J. Reprod. Fertil.* 25, 1-15.
- Hebel, R., and Stanberg, M. W. (1986). *Anatomy and Embryology of the Laboratory Rat*. Biomed Verlag, Worthsee.
- Kaufman, M. H. (1992). Assessment of developmental stage of pre and postimplantation mouse embryos. In *Atlas of Mouse Development*, Academic Press, San Diego.
- Krehbiel, R. H. (1937). Cytological studies of the decidual reaction in the rat during early pregnancy and in the production of deciduomata. *Physiol. Zool.* 10, 212-234.
- Lundkvist, O., and Ljungkvist, L. (1977). Morphology of rat endometrial stroma at the appearance of the Pontamine blue reaction during implantation after experimental delay. *Cell Tissue Res.* 184, 463-466.
- Lutwak-Mann (1962). Glucose, lactic acid and bicarbonate in rabbit blastocyst fluid. *Nature* 193, 653-664.
- Nicholas, J. S. (1962). Experimental methods and rat embryos. In *The Rat in Laboratory Investigations* (E. Ferris and J. Griffith, Eds.), Hafner, New York.
- Pincus, G. (1936). The eggs of mammals. *Exp. Biol. Monogr.*
- Psychoyos, A. (1971). Methods for studying changes in capillary permeability of the rat endometrium. In *Methods in Mammalian Methodology* (J. C. Daniel, Ed.), pp. 334-338. Freeman, San Francisco.
- Wimsett, W. A. (1975). Some comparative aspects of implantation. *Biol. Reprod.* 12, 1-40.

Handbook of Teratology

Edited by

JAMES G. WILSON

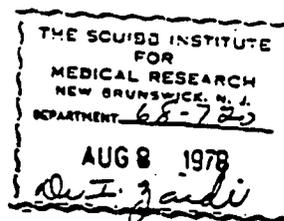
*The Children's Hospital Research Foundation
and Department of Pediatrics, University of Cincinnati
Cincinnati, Ohio*

and

F. CLARKE FRASER

*Departments of Biology and Pediatrics
McGill University and The Montreal Children's Hospital
Montreal, Canada*

4 Research Procedures and Data Analysis



PLENUM PRESS · NEW YORK AND LONDON ·

Methods and Concepts of 14 Biometrics Applied to Teratology

D. W. GAYLOR

I. INTRODUCTION

Some procedures are described which are used frequently in the statistical analysis of data from teratological studies. For the most part, the discussion here focuses on those species that produce litters with more than one offspring. There is a question of what teratological indices provide the most useful information. A discussion is presented on the proper choice of the experimental unit, the litter or fetus, for determining sample sizes. Owing to the large experimental variation often encountered in teratological studies, the desirability of replicated experiments is stressed. Measures of lethality and teratogenicity are discussed. Statistical tests for comparing treated and control animals are given. The use of fitting mathematical models to describe dose-response relationships is discussed.

II. MEASURES OF LETHALITY

A. Corpora Lutea, Implants, and Live Fetuses

If a compound is administered prior to nidation, the number of corpora lutea or number of implants may provide a measure of early lethality.

D. W. GAYLOR • National Center for Toxicological Research, Jefferson, Arkansas.

Frequently, compounds are not administered until after implantation, in which case only the number of live fetuses would serve as a measure of lethality.

Let c_j , m_j , and f_j be the number of corpora lutea, implants, and live fetuses for the j th litter, where $j = 1, 2, \dots, n$ is the number of pregnant mothers treated similarly. The best unbiased estimator of the average number of corpora lutea per litter is the simple arithmetic mean, \bar{c} , which is the sum of the corpora lutea for the n litters divided by the number of pregnant mothers, n :

$$\bar{c} = \frac{1}{n} \sum_{j=1}^n c_j = \frac{1}{n} (c_1 + c_2 + \dots + c_n)$$

The average number of implants and live fetuses per litter are computed in a similar manner.

The sample standard deviation for corpora lutea, s_c , is calculated by

$$s_c = \left[\frac{1}{n-1} \sum_{j=1}^n (c_j - \bar{c})^2 \right]^{1/2}$$

Similar calculations apply for the sample standard deviation for implants and live fetuses.

Even though the number of corpora lutea, implants, or live fetuses are not normally (Gaussian) distributed, the Central Limit Theorem states that means of nonnormal variables rapidly approach normality (Mood, 1950). Thus, approximate confidence limits on the mean number of corpora lutea can be calculated when $n \geq 10$ by $\bar{c} \pm t_c / \sqrt{n}$, where t is obtained from standard statistical t -tables with $(n-1)$ degrees of freedom for the level of confidence selected. This interval includes the true average number of corpora lutea, at the chosen level of confidence, that would be expected from all litters treated under similar conditions. Approximate confidence intervals for the average number of implants and average number of live fetuses per litter are calculated in a similar manner. The approximation is reasonably good for $n \geq 10$, but this does not mean that 10 is necessarily an adequate number of litters. The number of litters required depends upon the precision desired and the variance.

For example, suppose that the numbers of live fetuses in a group of 16 litters are: 9, 12, 10, 8, 7, 12, 11, 13, 8, 7, 11, 12, 10, 10, 12, and 8. The average litter size is 10 and

$$s_f = \left[\frac{1}{15} [(9-10)^2 + (12-10)^2 + \dots + (8-10)^2] \right]^{1/2} = 1.97$$

The 95% confidence interval for the mean number of live fetuses is $10.0 \pm 2.15(1.97)/\sqrt{16} = 10.0 \pm 1.05$.

The median sample size needed to achieve a given precision of $\pm d$ is given

by $n = t^2/d^2$, where t has the desired level of confidence and $n - 1$ degrees of freedom. The exact degrees of freedom, $n - 1$, cannot be determined until the sample size, n , is determined. As a first approximation the t -value with infinite degrees of freedom can be used. For example, if it is desired to determine the average number of live fetuses within ± 0.5 with 95% confidence, then $t = 1.96$. Using the estimate of the standard deviation for litter size of $s_r = 1.97$, the approximate median number of litters required is

$$n = \frac{(1.96)^2(1.97)^2}{(0.5)^2} = 59.6 \text{ or } 60 \text{ litters}$$

This procedure provides an estimate of sample size which is adequate to achieve the desired precision about half of the time.

To be more certain that the sample size is adequate, the required number of litters is calculated by

$$n = \frac{(t_1 + t_2)^2 s_r^2}{d^2}$$

where t_1 is the t -value with $n - 1$ degrees of freedom corresponding to the desired level of confidence, and t_2 is a one-sided t -value with $n - 1$ degrees of freedom corresponding to the probability that the sample size will be adequate to achieve the desired precision. As a first approximation, t -values with infinite degrees of freedom can be used. In the previous example, if a probability of 90% is desired that the sample size is adequate to estimate the average litter size within ± 0.5 with 95% confidence, then the estimated number of litters is

$$n = \frac{(1.96 + 1.28)^2(1.97)^2}{(0.5)^2} = 163 \text{ litters}$$

The discussion above on measures of lethality applies to the simple situation where a group of n pregnant animals is treated similarly under a given set of conditions at some point in time. The total number of animals available for experimentation at that time may be randomly assigned to different groups (e.g., different dosage levels, different routes of administration of a chemical). The estimates of the average responses and confidence limits apply only to the existing set of laboratory conditions.

If an experiment is replicated (repeated) at r different times or at r different laboratories, a broader base for inferences is provided. If \bar{f}_k represents the average for the mean number of fetuses per litter from the k th replicate ($k = 1, 2, \dots, r$), then an overall estimate of the average number of fetuses per litter is given by

$$\bar{f} = \frac{1}{r} \sum_{k=1}^r \bar{f}_k$$

This estimate of \bar{f} gives equal weight to each replicate. If the sample sizes, n_k , vary considerably among replicates and if the difference in the average responses among replicates is relatively small, it may be desirable to use an estimate of the average weighted by the number of litters per replicate. The calculation of the variance and confidence limits for these situations are considerably more complicated (Anderson and Crump, 1967).

The overall average number of corpora lutea or implants per litter from replicated experiments is treated in the same manner.

B. Percentages: Preimplantation Loss, Resorptions, and Dead Fetuses

The reason for using percentages is an attempt to obtain more precise measures of lethality by adjusting for variability in the number of corpora lutea or implants from litter to litter.

If a compound is administered early in gestation, the percentage of preimplantation loss may provide a measure of early lethality. This can be estimated by subtracting the number of implants from the number of corpora lutea, expressed as a percentage of all corpora lutea, or as the percentage of corpora lutea which result in implants.

The percentage of implants that resorb is an appropriate measure of embryoletality when a compound is administered after implantation occurs. However, this quantity may not be appropriate if a compound is administered before implantation, as an increase in the percentage of resorptions may result from a reduction in the number of implants.

If a distinction is made between late fetal death and resorptions (embryo death), then the percentage of the implants resulting in fetal deaths can be calculated. Again, this may not be an appropriate measure of fetolethality if a compound is administered prior to implantation. The total lethality is represented by the percentage of implants resulting in either embryoletality identified by resorptions or late fetal death or equivalently by the percentage of live fetuses.

For all percentages in this section, the average percentage is simply the total of the percentages divided by the number of litters. For example, the average percentage of resorptions in a group of n similarly treated mothers is

$$\bar{R} = \frac{1}{n} \sum_{j=1}^n R_j$$

where R_j is the percentage of resorptions for the j th litter.

The sample standard deviation for the percentage of resorptions is

$$s_R = \left[\frac{1}{n-1} \sum_{j=1}^n (R_j - \bar{R})^2 \right]^{1/2}$$

Where the number of litters is 10 or more, a reasonably good approximation for a confidence interval on the true percentage of resorptions is given by $R \pm t s_R / \sqrt{n}$, where t has the chosen level of significance and $n - 1$ degrees of freedom. Similarly, approximate confidence intervals can be calculated for the percentage of preimplantation loss or the percentage of dead fetuses. A situation that may arise, particularly for the percentage of dead fetuses, is that there may be no fetal deaths, giving $D_i = 0$ for all litters and $s_D = 0$. A confidence interval then could be based on the binomial distribution (see, e.g., Mood, 1950) for zero deaths out of $\sum n_i$ implants.

In an experiment replicated r times, an estimate of the overall average percentage of resorptions per litter is given by

$$\bar{R} = \frac{1}{r} \sum_{k=1}^r \bar{R}_k$$

where \bar{R}_k represents the average for the k th replicate. Alternatively, an estimate of the average weighted by the number of litters per replicate may be used. The variances of these estimators are discussed by Anderson and Crump (1967).

The overall average for the average percent of live fetuses or dead fetuses per litter from replicated experiments can be computed in a similar manner.

C. Percent of Litters with Resorptions or Dead Fetuses

In Section II.B the percentage of resorbed or dead fetuses was calculated for each litter and then averaged to obtain a measure of lethality. An additional measure is the percent of litters with resorbed or dead fetuses. Generally, if the percentage of resorptions per litter is high, it would be expected that the percentage of litters with resorptions would also be high. Thus, these two measures are usually highly correlated and reflect nearly the same thing—the extent of resorptions. However, it is possible that the average percentage of resorptions per litter could be quite high as a result of being concentrated in only a few litters. Thus, calculating the percent of litters with one or more resorptions provides a measure of whether resorptions are spread over litters or concentrated in a few litters; similarly, for dead fetuses.

Approximate confidence limits for the proportion of litters with resorbed or dead fetuses could be obtained from the binomial distribution. This is only a rough approximation, as a condition required for application of the binomial distribution is that each litter must have an equal probability of containing a resorbed or dead fetus. This condition undoubtedly is not satisfied, since litter sizes vary, which would result in differing probabilities of obtaining a resorbed or dead fetus.

III. MEASURES OF TERATOGENIC EFFECTS

A. Percent of Abnormal Fetuses

Let a_j represent the number of fetuses in the j th litter possessing a certain type of anomaly. This anomaly may be a specific type, such as cleft palate. Or, a_j may represent a group of anomalies, such as skeletal malformations. Or, a_j may represent the number of fetuses in the j th litter with any anomaly. In calculating the percentage of abnormal fetuses per litter, it is necessary to divide a_j by the number of fetuses, n_j , examined for that particular anomaly. For certain gross anomalies, all live fetuses are generally examined, so that $n_j = f_j$. However, for certain soft tissue anomalies, sectioning of organs may be required, or for skeletal anomalies special staining is required. It is common practice to select part of a litter for special staining for skeletal defects. Thus, the number of fetuses examined for a particular anomaly or group of anomalies is often less than the number of live fetuses in the litter. Thus, the percentage of animals with anomalies for the j th litter is calculated by

$$A_j = \frac{a_j}{n_j} \times 100\%$$

Unless an equal number of fetuses are examined for soft tissue anomalies and skeletal anomalies, the value of n_j will not be the same for each type of defect. The average percentage of anomalies per litter is

$$\bar{A} = \frac{1}{n} \sum_{j=1}^n A_j$$

where at least one fetus from each of n litters is examined for the anomaly. The sample standard deviation for the percentage of anomalies per litter is calculated by

$$s_A = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (A_j - \bar{A})^2}$$

For 10 or more litters, fairly good approximate confidence limits are given by $\bar{A} \pm t_A \sqrt{s_A/n}$, where t has the desired confidence level and $n - 1$ degrees of freedom. A situation that may arise is that there may be no anomalies of a particular type, giving $A_j = 0$ for all litters and $s_A = 0$. A confidence interval then could be based on the binomial distribution (see, e.g., Mood, 1950) for zero anomalies out of $\sum n_j$ fetuses examined.

In general, observing the percentage of abnormal fetuses per litter and the percentage of resorbed and dead fetuses per litter should give a sufficient indication of the fetotoxicity of a compound. However, it is quite possible that

there is a correlation between the percentage of abnormal fetuses and re-sorbed or dead fetuses. For example, the abnormal fetuses may tend to have a higher death rate. Thus, the percentage of abnormal fetuses per litter may not appear to be unusual. Therefore, it also is advisable to include the percentage of normal life fetuses per litter in any statistical analyses.

For experiments replicated r times, the overall average percentage of anomalies per litter is

$$\bar{A} = \frac{1}{r} \sum_{k=1}^r \bar{A}_k$$

where \bar{A}_k is the average number for the k th replicate. Again, weighted averages could be considered for percentages of abnormal fetuses (Anderson and Crump, 1967).

B. Percent of Litters with Abnormal Fetuses

In addition to calculating the average percentage of abnormal fetuses per litter, it is of interest to determine if anomalies occur throughout the litters or if they are concentrated in a few litters. This is accomplished simply by calculating the percentage of litters containing at least one abnormal fetus. This calculation can be based on a specific type of anomaly or can include all anomalies. Crude confidence limits could be obtained by applying the binomial distribution, but these limits are only rough approximations, as the probability of observing an anomaly in a litter changes, depending on the number of fetuses examined for that anomaly.

IV. FETAL WEIGHT

Whereas reduction in fetal weight due to treatment by a compound may be difficult to interpret biologically, fetal weight appears to be a very sensitive measure of toxicity. This may be due in part to the fact that weight is a continuous variable, as opposed to a discontinuous variable such as observing only whether or not a fetus is dead or possesses a particular anomaly. There is more statistical information in a continuous variable such as fetal weight. In general, it is advisable to keep male and female weights separate. If w_{ij} represents the fetal weight for the i th live male (or female) fetus in the j th litter, the best estimate of the average male fetal weight for the j th litter is

$$\bar{w}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} w_{ij}$$

where n_j is the number of male fetuses in the j th litter. If a litter does not contain any live male fetuses, that litter is not included in the calculation. If n

litters contain live male fetuses, an estimate of the average male fetal weight for that group of litters is

$$\bar{w} = \frac{1}{n} \sum_{j=1}^n \bar{w}_j$$

If an experiment is replicated r times, an estimate of the overall average male fetal weight is calculated by

$$\bar{w} = \frac{1}{r} \sum_{k=1}^r \bar{w}_k$$

where \bar{w}_k represents the average male fetal weight for the k th replicate. Again, weighted estimates could be considered (Anderson and Crump, 1967).

V. COMPARISON OF CONTROL AND TREATED GROUPS

A. Corpora Lutea, Implants, and Live Fetuses

A common objective of an experiment is to determine if a compound is lethal to embryos or fetuses or teratogenic when administered to laboratory animals under a particular set of conditions. A preliminary experiment may be needed to select a dosage regime which is not excessively lethal to the pregnant mothers. In a test comparing a treated group of animals with a control group of animals, the nonparametric Wilcoxon-Mann-Whitney U -test (see, e.g. Siegel, 1956) is recommended. This test makes no assumption as to the mathematical distribution of the response and still provides a fairly powerful test for determining if differences exist. When there are 10 or more litters per group, the ordinary t -test provides essentially the same results. A two-sided test is appropriate if one is interested in whether or not the compound causes either an increase or decrease in the response. In general, one would be interested in determining if a compound resulted in a decrease of corpora lutea, implants, or live fetuses; hence, a one-sided test would be utilized.

To illustrate the methods, consider the following results for the number of live fetuses obtained in 10 litters of mice treated with a compound compared with 10 control litters:

Control: 7, 8, 8, 10, 9, 5, 8, 7, 9, 9

Treated: 5, 8, 7, 7, 9, 8, 4, 7, 9, 6

The control group averages 8 live fetuses per litter and the treated group averages 7 live fetuses per litter. The question is: Does this 12.5% reduction in

litter size represent a statistically significant reduction in litter size due to the compound?

For the Wilcoxon-Mann-Whitney U -test, the data are ordered from lowest to highest values.

Control: 5, 7, 7, 8, 8, 8, 9, 9, 9, 10
Treated: 4, 5, 6, 7, 7, 7, 8, 8, 9, 9

The ordered rankings are

Control: 2.5, 7, 7, 12, 12, 12, 17, 17, 17, 20
Treated: 1, 2.5, 4, 7, 7, 7, 12, 12, 17, 17

where the average rank is assigned to tied values. The sum of the ranks for the control group is 123.5 and 86.5 for the treated group.

The value of U is determined from

$$U_c = n_c n_t + \frac{n_c(n_c + 1)}{2} - R_c$$

or

$$U_t = n_c n_t + \frac{n_t(n_t + 1)}{2} - R_t$$

where n_c and R_c are the sample size and sum of the ranks for the control group and n_t and R_t for the treated group. In this example,

$$U_c = 10 \times 10 + \frac{10 \times 11}{2} - 123.5 = 31.5$$

or

$$U_t = 10 \times 10 + \frac{10 \times 11}{2} - 86.5 = 68.5$$

In determining the level of significance for a comparison of two groups, the smaller value of U_c or U_t is used:

$$U = \text{minimum of } U_c \text{ or } U_t$$

Critical values of U are given by Siegel (1956) to obtain the level of significance. The value of 31.5 indicates a one-tailed significance level of approximately $P < 0.10$. That is, there is a 10% chance that the observed difference is not real but is due only to normal experimental variation.

For the parametric t -test of the same hypothesis, the variance among the control group is

$$s^2 = \frac{1}{9}[(7-8)^2 + (8-8)^2 + \dots + (9-8)^2] = 2.00$$

and for the treated group is

$$s_t^2 = \frac{1}{9}[(5-7)^2 + (8-7)^2 + \dots + (6-7)^2] = 2.67$$

An F -test with 9 and 9 degrees of freedom, $F = 2.67/2.00 = 1.33$, indicates that these two variances are not significantly different. Thus, a pooled variance with 18 degrees of freedom can be used for the t -test:

$$s^2 = \frac{18 + 24}{9 + 9} = 2.33$$

The t -test gives

$$t = \frac{8-7}{\sqrt{2.33(1+1)}} = 1.46$$

which is significant at $P < 0.10$ for a one-sided test. Thus, the result may have arisen by chance alone with a probability of about 10%. If the variances of the treated and control groups are unequal, an approximate t -test can be employed (see, e.g., Anderson and Bancroft, 1952).

One might ask the question: How many litters would be required on the average in order to demonstrate if a difference of $d = 1$ fetus per litter is significant at the 95% confidence level? The formula for the number of litters, n , for each group is

$$n = \frac{2t^2s^2}{d^2}$$

where as a first approximation, t is used with infinite degrees of freedom which would be $t = 1.64$ for a one-sided test. Thus...

$$n = \frac{2(1.64)^2(2.33)}{(1)^2}$$

or 13 litters in each group.

In order to determine the number of litters, n , required to have a high probability, P , of being able to detect a difference of d , use

$$n = \frac{2(t_1 + t_2)^2s^2}{d^2}$$

where t_1 is the one-sided or two-sided t -value corresponding to the significance level for the test and t_2 is the one-sided t -value corresponding to a

significance level of $1 - P$. As a first approximation, t -values with infinite degrees of freedom (normal deviates) can be used.

A one-sided test is used where it only is of interest to establish if the treated group exhibits a change in one direction from the control group. A two-sided test is used where it is of interest to determine if there is a change in either direction, increase or decrease, from the controls. For example, the two-sided normal deviate corresponding to the 5% significance level is 1.96 and the one-sided normal deviate is 1.64.

In comparing data based on counts the square root of the counts are sometimes used to stabilize variances.

B. Percentages: Preimplantation Loss, Resorptions, Dead and Abnormal Fetuses

For comparing percentages, some investigators have combined all the data from all litters and calculated the percentage of abnormal fetuses in the control groups and treated groups. The percentages of fetuses responding in the two groups have then been compared by the common 2×2 chi-square test. The procedure implies that each conceptus is an independent experimental unit with an equal probability of being defective. That is, it assumes that each conceptus in a litter or a conceptus in one litter is no more likely to receive more or less of the compound than a conceptus in another litter. Thus, pooling of data assumes no maternal effect. That is, it assumes that each pregnant mother handles a compound in a like manner, such that metabolism, excretion rates, and so on, are similar to the extent that the offspring can be thought of as coming from one large homogeneous litter. Obviously, such an assumption seldom would be warranted. Since the compound is administered to the pregnant mother, the mother or litter is the experimental unit. This topic has been addressed by a number of authors (Weil, 1970; Healy, 1972; Kalter, 1974; Staples and Hase-man, 1974; Becker, 1974).

A recommended procedure is to compute P , R , D , T , and A ; the percentages of preimplantation loss, resorptions, dead, resorbed and dead, and abnormal fetuses, respectively, for each litter. Percentages between control and treated animals then can be compared by the Mann-Whitney-Wilcoxon U -test (see, e.g., Siegel, 1956) or generally for sample sizes greater than 10 by the t -test (see, e.g., Anderson and Bancroft, 1952). A transformation commonly used for the t -test with proportions, p , is to use the arcsin \sqrt{p} , which nearly stabilizes the variances if the denominators of the proportions are of nearly equal size.

To illustrate the procedures for comparing percentages, consider the following data on resorptions in mice:

Control Group

Litter (j)	Implants (n _j)	Resorptions (r _j)	Percent resorptions (R _j)
1	7	0	0
2	9	1	11.1
3	9	2	22.2
4	8	1	12.5
5	6	0	0
6	10	0	0
7	7	0	0
8	9	1	11.1
			7.1 average

Treated Group

Litter (j)	Implants (n _j)	Resorptions (r _j)	Percent resorptions (R _j)
1	8	1	12.5
2	8	0	0
3	9	1	11.1
4	6	2	33.3
5	8	3	37.5
6	9	0	0
			15.7 average

For this example, $U = 16.5$, which is significant at approximately $P < 0.19$. There is not strong statistical evidence that the increased percentage of resorptions was due to the treatment, as there is approximately a probability of 0.19 that this difference is due to chance alone.

The problem above serves to illustrate the weakness of using too few litters in teratological studies. Even though the percentage of resorptions more than doubled in the treated group, this experiment was unable to provide strong statistical evidence that the effect was real.

The formula in the previous section could be used to determine approximately the number of litters required. In this case, using the transformation $\arcsin \sqrt{p}$, where p is the observed proportion, gives $s^2 = 0.0598$ for the example given above. If a one-sided t -test at the 5% significance level is used and it is desired to detect a doubling of resorptions from 7% in the controls to 14% in the treated group with a probability of at least 0.80, then the number of litters per group should be approximately

$$n = \frac{2(1.64 + 0.84)^2(0.0598)}{(0.384 - 0.268)^2} = 55 \text{ litters}$$

It is important to remember that statistical tests are only tools to be used by the scientist to interpret data. The quantities used to indicate toxicity may be correlated, requiring caution of interpretation. For example, the number of resorptions, dead, and abnormal fetuses may be correlated. The number of abnormal fetuses may decrease if there is a tendency for abnormal fetuses to be resorbed. In addition to the percentages calculated above, it is advisable to include the percentages of normal fetuses per litter in any analyses. Since the objective of most teratological studies is to look for adverse effects, there is a tendency to ignore measures of normalcy.

C. Percent of Litters with Resorptions and Dead and Abnormal Fetuses

The techniques discussed in the previous section probably offer the best methods for detecting toxic effects. However, it is possible that the average percent of resorbed, dead, or abnormal fetuses may be quite high, simply because of a concentration of these conditions in only a few litters. For example, particularly small litters may have a high probability of being almost entirely resorbed, dead, or abnormal. Thus, a few litters could raise the average response considerably. Therefore, it is necessary to consider the distribution of a defect across litters. This can be accomplished by comparing controls and treated groups for the percentage of litters possessing a particular defect. A discussion of comparing two percentages is given by Snedecor and Cochran (1967). If the total number of litters is greater than 40, a chi-square test corrected for continuity of a 2×2 contingency table is suggested. The data may be displayed by a 2×2 table.

	Number of litters		
	No defects	One or more defects	Total
Controls	a	b	n_c
Treated	c	d	n_t
Total	a + c	b + d	n

The chi-square value is calculated by

$$\chi^2 = \frac{(ad - bc) - n/2)^2 n}{n_c \times n_t \times (a + c) \times (b + d)}$$

This value can then be compared to the values tabulated in a chi-square table with one degree of freedom to determine the level of significance of the difference between b/n_c and d/n_t . For a one-sided test, compare $\sqrt{\chi^2}$ with a standard normal deviate.

For example, suppose that 4 out of 20 (20%) control litters contained abnormal fetuses and 12 out of 30 (40%) of the treated litters contained one or more abnormal fetuses. Thus, $a = 16$, $b = 4$, $c = 18$, $d = 12$, $n_c = 20$, $n_t = 30$, and $n = 50$. Then,

$$\chi^2 = \frac{(16 \times 12 - 4 \times 18) - 25)^2 \times 50}{20 \times 30 \times 34 \times 16} = 1.3825$$

giving $\sqrt{\chi^2} = 1.18$, which is significant at approximately $P < 0.12$ for a one-sided test.

Again, the question can be raised: How many litters would be required to have a high probability of detecting doubling of the percentage of litters containing abnormal fetuses from 20% of the control litters to 40% of the treated litters? The approximate number of litters required, n , in each group in order to have a high probability of detecting a specified difference is given by

$$n = \frac{(Z_1 + Z_2)^2}{2(\arcsin \sqrt{P_t} - \arcsin \sqrt{P_c})^2}$$

where P_t and P_c are the expected proportions in the treated and control groups, Z_1 is the normal deviate corresponding to the significance level of a one- or two-tailed test, and Z_2 is the one-sided normal deviate corresponding to $1 - P$, where P is the minimal desired probability of detecting the difference between P_t and P_c . In this example, if it is desired to detect a difference between 20% and 40% with a probability of 0.8 using a one-sided significance level of 0.05,

$$n = \frac{(1.64 + 0.84)^2}{2(\arcsin \sqrt{0.4} - \arcsin \sqrt{0.2})^2} = 63 \text{ litters per group}$$

A word of caution is necessary in comparing the proportion of litters exhibiting a certain defect. The chi-square test is only approximate because it assumes that each litter has the same probability of containing a defect. Since litter sizes vary, this condition is not satisfied. For a minimal condition for using this test, it is necessary that the average number of implants per litter is approximately equal for the control and treated groups. This should pose no problem for compounds that are administered after implantation takes place during gestation. However, if the number of implants is different between the control and treated litters, the test may be meaningless, as the larger litters may have a higher probability of containing a resorbed, dead, or abnormal fetus.

D. Fetal Weight

The fetal weights for the two sexes should be analyzed separately. Fetal-weight data generate a three-level nested classification: control versus treated groups, litters within groups, and fetuses within litters. Since the number of male or female fetuses per litter will not be equal, the analysis of variance which provides an approximate *F*-test for the difference between groups is somewhat complicated and is not presented here (see, e.g., Anderson and Bancroft, 1952).

Fetal weight appears to be a sensitive indicator of toxicity. It generally is a more consistent measure than the numbers of implants, resorptions, etc., and percentages of dead fetuses, abnormal fetuses, etc. Even though the biological significance of a reduction in fetal weight may be difficult to assess, it generally is an indication of the presence of other toxic effects.

E. Replicated Experiments

The examples given generally indicate the need for more litters than are typically used in teratological experiments. Thus, a moderate number of litters may be required. Laboratory facilities and resources may be inadequate to accommodate a large number of litters simultaneously. This situation is easily resolved by conducting the experiment in blocks (replicates) over a period of time. Animals in each replicate would be impregnated on the same date and must be assigned at random to the treatment and control groups. For example, a replicate may consist of 10 treated and 10 control litters started on the experiment on a given date. This process is repeated at later dates until an adequate number of litters are obtained.

The replicated experiment has an inherent advantage over a study conducted at one time. Considerable differences in responses have been noted between groups of animals treated at the same laboratory at different times even though laboratory conditions were supposedly similar. Thus, an experiment conducted at only one time may lead to different conclusions than the same experiment conducted at a different time. Reasons for these differences may be many and unknown and unfortunately can only be attributed to the rather vague but well-known existence of experimental variation. The replicated experiment provides an opportunity to provide a broader base for inferences from the data, as the average results are likely to be more representative and reproducible since the data are averaged over a wider set of conditions. One does not lose precision in such an experiment, because the treated and control groups are still compared within a replicate under similar conditions.

The statistical analyses of replicated experiments become more complex by extending the *U*-test to nonparametric analysis-of-variance techniques (see,

e.g., Siegel, 1956) or extending the *t*-test to parametric analysis-of-variance techniques (see, e.g., Anderson and Bancroft, 1952). Experiments involving more than one strain, route of administration, or days of treatment during gestation again require more complicated statistical analyses likely to include analysis-of-variance techniques for comparing several groups simultaneously. Snedecor and Cochran (1967) provide a discussion for combining results from a series of 2×2 tables for comparing percentages.

VI. MULTIPLE-DOSE EXPERIMENTS: DOSE-RESPONSE

It is common to utilize several dosages in a biological experiment in order to study the intensity of a response as a function of dose. The establishment of a dose-response curve provides strong evidence of a cause-and-effect relationship between the administration of a compound and the observed biological response. A dose-response curve also may provide some indication as to dosage levels that may only produce negligible biological effects.

There are a multitude of types of dose-response regression curves which may be fit to data. No attempt will be made here to discuss the techniques available. Many statistical texts discuss curve fitting. In biology, particular attention has been given to fitting curves to quantal (percentage) data by probit analysis (Finney, 1971).

REFERENCES

- Anderson, R. L., and Bancroft, T. A., 1952 *Statistical Theory in Research*, McGraw-Hill, New York.
- Anderson, R. L., and Crump, P. P., 1967, Comparisons of designs and estimation procedures for estimating parameters in a two-stage nested process, *Technometrics* 9:499-516.
- Becker, B. A., 1974, The statistics of teratology, *Teratology* 9:261-262.
- Finney, D. J., 1971, *Probit Analysis*, 3rd ed., Cambridge Univ. Press, London.
- Healy, M. J. R., 1972, Animal litters as experimental units, *J. Roy. Stat. Soc., Appl. Stat.* C21:155-169.
- Kalter, H., 1974, Choice of the number of sampling units in teratology, *Teratology* 9:257-259.
- Mood, A. M., 1950, *Introduction to the Theory of Statistics*, McGraw-Hill, New York.
- Siegel, S., 1956, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, New York.
- Snedecor, G. W., and Cochran, W. G., 1967, *Statistical Methods*, Iowa State Univ. Press, Ames, Iowa.
- Scaples, R. E., and Haseman, J. K., 1974, Selection of appropriate experimental units in teratology, *Teratology* 9:259-60.
- Well, C. S., 1970, Selection of the valid number of sampling units and a consideration of their combination in toxicological studies involving reproduction, teratogenesis, and carcinogenesis, *Food Cosmet. Toxicol.* 8:177-182.

16 Page(s) Redacted

Draft

Labeling