Reviewer Guidance

Integration of Study Results to Assess Concerns about Human Reproductive and Developmental Toxicities

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 120 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to Dockets Management Branch (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact Joseph J. DeGeorge, 301-594-5476.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
October 2001

Pharmacology/Toxicology

Reviewer Guidance

Integration of Study Results to Assess Concerns about Human Reproductive and Developmental Toxicities

Additional copies are available from:

Drug Information Branch (HFD-210)
Center for Drug Evaluation and Research (CDER)
5600 Fishers Lane, Rockville, MD 20857, (Tel) 301-827-4570
Internet at http://www.fda.gov/cder/guidance/index.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
October 2001

Pharmacology/Toxicology

TABLE OF CONTENTS

Reviewer Guidance[1]

Integration of Study Results to Assess Concerns about Human Reproductive an

Developmental Toxicities

This draft guidance, when finalized, will represent the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statutes and regulations.

I. INTRODUCTION

This draft guidance describes a process for estimating the increase in human developmental and reproductive risks as a result of drug exposure when definitive human data are unavailable. The overall approach integrates nonclinical information from a variety of sources (i.e., reproductive toxicology, general toxicology, and toxicokinetic and pharmacokinetic information, including absorption, distribution, metabolism and elimination findings) and available clinical information to evaluate a drug's potential to increase the risk of an adverse developmental or reproductive outcome in humans.

The integration process focuses on the likelihood a drug will increase the risk of adverse human developmental or reproductive effects. It does not consider the nature (e.g., severity, reversibility or repairability) of the adverse response, or otherwise consider the clinical implications of the response. These risk management issues will be discussed in separate guidance on how to address the clinical implications of developmental and reproductive risks in product labeling. Because of inherent differences between drug and biological products, and resulting differences in the types of preclinical data collected for drug and biological products, the process described in this guidance will often not be useful in evaluating potential adverse reproductive or developmental effects for

[1] This guidance has been prepared by the Office of Review Management in the Center for Drug Evaluation and Research (CDER) at the Food and Drug Administration (FDA).

biological products. However, the general principles described (i.e., a comprehensive analysis of available data) will typically be of some relevance to biological products.[2]

II. BACKGROUND

A. Data Needed to Use the Integration Process

Ordinarily, the integration process should be based on an evaluation of a complete set of the expected general toxicology, reproductive toxicology, and pharmacokinetics studies.[3] This evaluation should include an assessment of the ability of the drug to produce a positive finding in the relevant animal studies (e.g., whether doses used were large enough to induce toxicity of some kind). The evaluation should also compare animal and human pharmacodynamic effects, animal and human metabolism and disposition, animal and human pharmacologic and toxic effects, and drug exposures in animal studies in relation to the highest proposed dose in humans.

The type and extent of available toxicology data may vary depending on the biological actions of the product, test systems available for studying the compound, and other factors. In some cases, the data will not include all desirable general toxicology, reproductive toxicology, and pharmacokinetics studies. In some of those cases, it may still be possible to use the integration process without all the desired information. In other cases, limited available data will preclude the use of the integration process (e.g., often the case for biological products). Even if the integration process cannot be used, the product should be evaluated to the greatest extent possible in accordance with sound scientific principles and the considerations described in this document.

B. Types of Reproductive and Developmental Toxicity Evaluated

For purposes of this document, there are two broad categories of toxicity — reproductive and developmental toxicity — and, within those categories, seven classes of toxicity. In the reproductive toxicity category there are three classes of toxicity: toxic effects on *fertility, parturition*, and *lactation*. In the developmental toxicity category there are four classes of toxicity: *mortality, dysmorphogenesis (structural alterations), alterations to growth*, and *functional toxicities.* For a given drug, each class of toxicity should ordinarily be assessed. A positive signal in any class of reproductive or developmental toxicity, whether in valid

[2] Although this is not a joint CDER/CBER guidance, CBER was consulted during guidance development. For more information, contact the Division of Clinical Trials Design and Analysis.

[3] See the following International Conference on Harmonisation (ICH) guidances for industry: *M3 Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals; S3A Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies; S5A Detection of Toxicity to Reproduction for Medicinal Products;* and *S5B Detection of Toxicity to Reproduction for Medicinal Products: Addendum on Toxicity to Male Fertility*.

reproductive or general toxicology studies or from human use studies, should be evaluated using the process described in this guidance to estimate the likelihood of increased reproductive or developmental risk for humans (see discussion of the integration process in Section III and schematic representation of the process in Figure C).

1. Reproductive Toxicities

Reproductive toxicities include structural and functional alterations that may affect reproductive competence in the F_0 generation. The three classes of reproductive toxicity include effects on fertility, parturition, and lactation.

- Fertility

Male reproductive toxicity associated with administration of a drug may be seen as degeneration or necrosis of the reproductive organs, reduction in sperm count, alterations to sperm motility or morphology, aberrant mating behavior, altered ability to mate, alterations to endocrine function, or overall reduction in fertility.

Female reproductive toxicity may be seen as damage to the reproductive organs, alterations to endocrine regulation of gamete maturation and release, aberrant mating behavior, altered ability to mate, or overall reduction in fertility. Diminished fertility in female animals is typically detected by reductions in the fertility index, the number of implantation sites, time to mating, or fecundity.

- Parturition

Toxicities affecting labor and delivery in animals may be seen as changes in the onset or duration of parturition. Changes in the duration of parturition are frequently reported as mean time elapsed per pup, or total duration of parturition.

- Lactation

Drugs administered to lactating animals may be a source of unwanted exposure in the nursing neonate, may alter the process of lactation in the nursing mother (e.g., the quality or quantity of milk), or may alter maternal behavior towards the nursing offspring.

2. Developmental Toxicities

Developmental toxicities are generally those that affect the F_1 generation. The four classes of developmental toxicity are mortality,

dysmorphogenesis (structural alterations), alterations to growth, and functional toxicities.

- Mortality

Mortality due to developmental toxicity may occur at any time from early conception to post-weaning, ("embryo-fetal death" is a subset of mortality due to developmental toxicity). Thus, a positive signal may appear as pre- or peri-implantation loss, early or late resorption, abortion, stillbirth, neonatal death or peri-weaning loss.

- Dysmorphogenesis (Structural alterations**)**

Dysmorphogenic effects are generally seen as malformations or variations to the skeleton or soft tissues of the offspring, and are commonly referred to as *structural alterations*.

- Alterations to Growth

Alterations to growth are generally seen as growth retardation, although excessive growth or early maturation may also be considered alterations to growth. Body weight is the most common measurement for assessing growth rate. Crown-rump length, and ano-genital distance may also be measured.

- Functional Toxicitie**s**

Functional toxicities could include any persistent alteration of normal physiologic or biochemical function, but typically only developmental neurobehavioral effects and reproductive function are measured. Common assessments include locomotor activity, learning and memory, reflex development, time to sexual maturation, mating behavior, and fertility.

III. DISCUSSION

The complete data integration process is schematically presented in Figures A-C, which are attached to this document. To clarify the manner in which data should be passed through the integration process, the process has been divided into three components, which are discussed in the following sections A-C. Briefly, Figure A is applicable to all data-sets, while Figure B is applicable only to data-sets without evidence of reproductive or developmental toxicity. Figure C is applicable to data-sets with positive indications of reproductive or developmental toxicity.

A. Overall Decision Tree (Figure A).

The decision tree process outlined in Figure A at the end of the document, should be used ***for each of the seven classes of reproductive or developmental toxicity*** discussed in Section II.B. For a given drug, studies may have been conducted to evaluate potential effects on none, some, or all of the classes of reproductive and developmental toxicity. Where studies are available for any of the different classes, the outcome may be one or more positive signals, or no signal. It is recognized that in practice one study may address several classes of toxicity and that a study may be considered adequate to evaluate all, some, or none of the classes of toxicity addressed. Figure A depicts the sequential decisions that should be made in evaluating the various situations that may be encountered and the next steps that should be taken where there are evaluable studies with positive or negative findings.

1. Availability of Studies

In Figure A, the first question that should be asked for each class of toxicity is: "Were studies performed that are relevant to an assessment of the risk of that class of reproductive or developmental toxicity in humans and are the detailed study results available for comprehensive evaluation?"

If no studies were conducted, or detailed study results are unavailable for comprehensive evaluation, the review should explain that studies adequate to assess the risk of that class of toxicity were not done, or are otherwise unavailable. In such circumstances, risk to humans is considered *unknown* or *not evaluable* and the product labeling should reflect that conclusion:

> **Example: The risk of [class of reproductive or developmental toxicity] with [Drug X] is unknown. There are no data to evaluate its potential to cause [class of reproductive or developmental toxicity].**

If studies were conducted and are available for comprehensive evaluation, the assessment process should continue with question 2.

2. Relevance of Studies

The next question that should be asked for each class of toxicity is: "Do the studies done provide information relevant to assessing the risk of that type of reproductive or developmental toxicity for the proposed human use?" If the test system was not relevant, the review should explain why the studies were not relevant or otherwise appropriate (i.e., inappropriate test protocol or species, nonrelevant route of drug administration[4]) and

[4] This may also apply to information from humans in which the route is inappropriate to provide relevant information for the clinical indication.

should discuss all supporting information that bears on study relevance. If the test system was not relevant, the risk to humans is considered *unknown* or *not evaluable* and the product labeling should reflect that conclusion:

> **Example: The risk of [class of reproductive or developmental toxicity] with [Drug X] is unknown. There are insufficient data to evaluate its potential to cause [class of reproductive or developmental toxicity].**

If the studies conducted are relevant to evaluating the risk of the particular class of toxicity in humans, the risk integration process should continue with question 3. Note that the processes in Figures B and C (see end of document) are intended to be used only when studies are considered adequate to assess the specified risk. They should not be used to evaluate findings (positive or negative) derived from inadequate studies.

3. Presence or Absence of a Signal

If the test system is relevant and appropriate for assessing the risk of toxicity in humans, the next question that should be asked for each class of toxicity is "Was there a positive signal (suggesting toxicity)?" If no signal was seen, the evaluation process should continue per Section B (Figure B at end of document). If a positive signal was seen, the evaluation process should continue per Section C (Figure C at end of document).

B. No Signal (Figure B)

Where there is no positive signal for one of the seven classes of reproductive or development toxicity, the risk assessment should be a step-wise process leading to a recommendation about the relevance of the nonfinding in humans. A graphic representation of this process is presented in Figure B (see end of document).

If multiple studies are available to assess a class of reproductive or developmental toxicity (e.g., multiple studies would be expected for the evaluation of dysmorphogenic effects - ICH stage C), the process in Figure B should be used only if the results of all studies relevant to a particular class of reproductive or developmental toxicity are negative for that type of toxicity. If any study (general toxicity, reproductive, or developmental toxicology study) has a positive signal for that class of reproductive or developmental toxicity, the process in Section C (Figure C) should be used.[5]

The following four factors should be considered during the evaluation of each class of reproductive or developmental toxicity for which there was no signal.

[5] Studies with conflicting signals and inter-study *concordance* and *nonconcordance* are addressed in Section C.

1. The Model/Test Species Predictive Adequacy

To what extent are the models or test species used likely to be predictive of human response? The following questions bear on the determination of a model's predictive adequacy.

- Do any of the models or test species (or systems) demonstrate or have the capability of responding to the pharmacodynamic effect(s) of the drug?

- Do any of the model/test species (or systems) demonstrate an overall toxicity profile that is relevant to the human toxicity profile?

- Do any of the model/test species (or systems) demonstrate pharmacokinetic (including ADME) profiles for the drug that are qualitatively similar to those in humans?

If the responses to these questions suggest that the response of the test species is of little relevance to humans, the review should explain why the animal study or studies conducted with the drug may not be fully adequate to evaluate the risk for the particular class of toxicity in humans (i.e., why the test may have low predictive value). Even if the test system is determined to be of limited relevance, the review should consider the remaining factors (2-4 below) and describe any additional uncertainties.

2. Adequacy of Study Doses and Exposure

The following elements should be considered in assessing the relevance to humans of the drug doses and exposure in the test system:

- Were adequate doses (concentrations) of the drug administered to the test species or test systems (e.g., MTD, MFD)?

- Were the exposures (based on AUC, C_{max}, or other appropriate systemic exposure metric) achieved in the test species or test systems adequate relative to those demonstrated in humans at the maximum recommended human dose?

If the answer to either of these questions is no, the evaluation should state that the animal studies conducted may be inadequate to fully evaluate the risk for the particular class of toxicity reported to be negative and explain in detail why they may be inadequate. Even if the study doses and exposure are considered inadequate, the evaluation should proceed to the remaining sections (3-4 below), and any additional uncertainties should be described.

3. Class Alert

Class alerts should be based on adverse reproductive or developmental effects previously demonstrated in humans by closely related chemical entities or compounds with similar pharmacodynamic effects. If there is a Class Alert for the drug, based on a related chemical structure of parent or metabolite or related pharmacologic effect, the class-specific information relevant to the class of toxicity reported to be negative should be included in the risk evaluation and discussion of the drug. The basis for the class alert for adverse effects in humans should be reasonably applicable to the drug being evaluated.

4. Signals for Related Types of Reproductive and Developmental Toxicity

The next step in evaluating the relevance of a no signal finding for a particular class of reproductive or developmental toxicity is to assess findings for related reproductive and developmental toxicities. A positive signal for one class of toxicity may suggest some risk in humans for other toxicities in the same category for which there were no findings in animals. The issue of related toxicities is most relevant for developmental toxicities. For example, if there is no signal for fetal mortality, but a positive signal for alterations to growth or dysmorphogenesis in one (or more) animal species, it may be inappropriate to conclude there is no risk of fetal mortality for humans. Related toxicities may also be relevant for reproductive toxicities where a hormonal mechanism is identified, the mechanism could be relevant to multiple aspects of reproductive performance, and the mechanism is relevant to humans.

If positive signals for related classes of toxicity were observed in the animal studies, the evaluation should state that there was no observed effect on the type of toxicity being assessed, but positive signals were seen for related toxicities. If there is no positive signal for any class of reproductive or developmental toxicity, the evaluation should state that there is no expected increase in risk for reproductive or developmental toxicity in humans based on the results of animal studies.

C. One or More Positive Signals (Figure C)

1. Overview of the Integrative Process

There are six factors that may affect the level of concern for a positive signal in any of the classes of reproductive or developmental toxicity: (1) signal strength part I, (2) signal strength part II, (3) pharmacodynamics, (4) concordance (metabolic and toxicologic concordance to the human); (5) relative exposure; and (6) class alerts. As described in more detail below, the integration tool considers signal strength in two different ways,

so signal strength is treated as two separate factors. Each factor has several contributory elements. The outcomes of the analyses of these six factors are used in the six columns in the integration tool (see Figure C). Human data may be considered separately from nonclinical findings and may greatly influence the overall assessment of human risk of reproductive or developmental toxicity.

The overall integrative analysis begins with a positive signal in a class of reproductive or developmental toxicity in one or more of the examined species. The positive signal may be from a reproductive or developmental toxicology study or an effect observed on a reproductive tissue, system, or behavior in a general toxicology study. Each of the six factors should be analyzed independently. Guidance is provided on what types of observations for each of the six factors might increase, decrease, or leave unchanged the level of concern for that factor. These analyses should not be an arithmetic summation of the contributing elements within each factor, but a weighted integration that takes into account the quality and nature of the data under consideration. The assessments of concern for each of the six factors should be assigned values of +1, -1 or 0, respectively, if the factor is perceived as increasing, decreasing, or leaving unchanged the level of concern for a class of reproductive or developmental toxicity. Conclusions from the six analyses should be summed to arrive at a comprehensive evaluation of the potential increase in risk for each class of the seven reproductive or developmental classes for which there was a positive signal.

2. A Note on Intra- and Inter-Species Concordance

Intra- and interspecies concordance of adverse effects in animals deserves some special consideration in this risk integration process. Positive signals in related types of reproductive or developmental toxicity within the same species indicates intra-species concordance of effects (e.g., a reduction in normal growth and an increase in developmental mortality). Positive signals for the same or a related type of toxicity across species indicates interspecies concordance. In general, findings for which there is intra- or interspecies concordance are more convincing than a positive signal in only one toxicity class in only a single species.

In evaluating potential human risk for adverse reproductive or developmental outcomes, if there is interspecies concordance for a single adverse effect it may be reasonable to conclude that a similar effect is the most likely adverse event to be seen in humans treated with the drug. If different but related adverse effects are seen in multiple test species (e.g., alterations to growth in one species and developmental mortality in another, or parturition effects in one species with lactation effects in the second), it may be reasonable to assume there is some level of risk for categorically related endpoints in humans.

A detailed discussion of the overall integrative analysis, the six individual factors, the contributory elements for each factor, and the assignment of the level of concern for each factor is presented in Sections 1-6**.**

a. Signal Strength, Part I

For the first signal strength factor, a positive signal in any reproductive or developmental toxicity class should be analyzed with respect to three contributory elements that examine whether the finding is present in more than one setting: (a) whether there is cross-species concordance (where more than one species has been studied), (b) whether there is multiplicity of effects, and (c) whether adverse effects are seen at more than one time.

Cross-Species Concordance

The defining characteristic of cross-species concordance is a positive signal in the same class of reproductive or developmental toxicity in more than one species. Cross-species concordance is most likely to be identified for structural abnormalities (dysmorphogenesis) or developmental mortality because these toxicities are frequently detected in the *organogenesis* testing paradigm, in which multiple species are typically evaluated. In addition, alterations to endocrine function or gonadal histopathology (which may alter fertility) may be indirectly detected in subchronic and chronic toxicity studies in rodents and nonrodents. When cross-species concordance is observed, there is increased concern for reproductive or developmental toxicity in humans. In contrast, there is decreased concern when a signal is detected in only one species (with the proviso that the negative species is an appropriate animal model and the studies were adequate in design, dosing, and implementation).

For alterations to parturition or lactation, it's often not possible to assess cross-species concordance because peri- and postnatal studies to assess these classes of toxicity are usually done in only a single species.

Multiplicity of Effects

Multiplicity of effects refers to observation, in a single species or animal model, of two or more positive signals within one of the two general categories of toxicity (reproductive or developmental) or within one of the seven classes of reproductive or developmental

toxicities. The observation of increased embryo-fetal death and structural abnormalities (dysmorphogenesis) in an animal test species is an example of multiple positive signals within a general category. The observation of two or more positive signals for structural abnormalities in tissues of multiple embryonic origin (e.g., defects affecting soft tissue, skeletal tissue, and/or neural tissue) is an example of multiple positive signals in a toxicity class.

If all species examined demonstrate multiplicity of effects, there is increased concern for reproductive or developmental toxicity in humans. If there are positive signals in two or more species, but multiplicity of effects is observed in only one species, concern is unchanged for this element. If no species studied exhibits multiplicity of effects, there is decreased concern.

Adverse Effects at Different Stages of the Reproductive or Developmental Process

Evidence of toxicity may arise during any stage of the reproductive or development process. For example, developmental mortality may be reported as early or late resorptions, abortions, or stillbirths. If a positive signal in animals is observed in multiple stages of development, there is generally greater concern for adverse human reproductive outcomes. If a positive signal is observed only during a single, discreet interval, the level of concern is unchanged. If the positive signal occurs only during processes that are of limited relevance to humans (rare), there would be less concern for adverse human reproductive outcomes. In addition to its relevance to this evaluation process, it is also important to define the timing of the period of susceptibility for the observed positive signal to provide a context for the human risk.

b. Signal Strength, Part II

In assessing the second signal strength factor, a positive signal should be analyzed with respect to the following three contributory elements: (a) co-existence of maternal toxicity, (b) presence of a dose-response relationship, and (c) the observation of rare events.

Maternal Toxicity

In weighing a signal of toxicity, the magnitude of adverse effects in the offspring versus the severity of maternal (and, for fertility studies, paternal) toxicity should be considered when drawing a conclusion about the relevance of the F_0 toxicity to effects observed in the offspring. This assessment is relevant to all seven classes of reproductive and developmental toxicity. A positive signal occurring

at doses that are not maternally toxic increases concern for human reproductive or developmental toxicity. If a positive signal is observed only in the presence of frank maternal toxicity, there is decreased concern, provided that the positive signal may be reasonably attributed to maternal toxicity.[6]

When evaluating a positive signal in two or more species, assessment of the implications of maternal or paternal toxicity should be based on a composite analysis of the data from all adequately studied species. If a positive signal is seen in two or more species in the absence of maternal toxicity, there is increased concern for adverse human reproductive outcomes. If a positive signal is seen only in the presence of clear relevant maternal toxicity in multiple species, there is decreased concern. If there is nonconcordance between test species as to the presence and relevance of maternal toxicity, there may be no change in the overall level of concern for this contributory element.

If any species is considered inappropriate to assessing the implications of maternal or paternal toxicity, the evaluation should be performed using the remaining available data.

Dose-Response Relationship

Concern for human reproductive or developmental toxicity is increased when a positive signal is characterized by any of the following: (1) increased severity of adverse effects with an increase in dose, (2) increased incidence of adverse effects with an increase in dose, or (3) a high incidence of adverse effects across all dosed groups. Conversely, the absence of all three of these indicia of dose-response would be cause for unchanged or decreased concern.

If multiple species are evaluated, a clear dose response across all tested species increases concern. If a positive signal occurs in more than one species, only one of which demonstrates one of the dose-response relationships described above, the level of concern will generally be unchanged. If there is no dose-response in any species, there is decreased concern for this contributory element.

Rare Events

[6] The attribution of adverse fetal effects to maternal (or paternal) toxicity can be based on previously collected data demonstrating the relationship between the maternal/paternal and reproductive effects.

Developmental toxicity studies usually lack the statistical power to detect subtle increases in rare events. Thus, an increased frequency of positive signals for rare events in drug-exposed animals increases concern for reproductive or developmental toxicity in humans. The absence of an increased frequency of rare events, however, does not decrease concern.

When multiple species (more than two) are studied, an increased frequency of positive signals for rare events in more than one species increases concern for adverse outcomes in humans even if not all species have an increased frequency of positive signals.

c. Pharmacodynamics

A positive signal should be analyzed with respect to the following three pharmacodynamic elements: (a) the therapeutic index, (b) biomarkers as a benchmark, and (c) the similarity between the pharmacologic and toxicologic mechanisms.

Therapeutic Index (TI)

The TI is used to identify the extent to which there is overlap between therapeutic doses and doses that cause reproductive or developmental toxicity. It is unusual to obtain well-defined dose-response curves for toxicity and efficacy from a single species. Thus, the use of estimations or surrogate endpoints (related to the therapeutic mechanism) for this evaluation may be warranted. To reduce the impact of variation in the slope of the dose-response curves, estimation of the TI should generally be based on comparison of the TD_{10} and the ED_{90} concentrations.[7]

If the $TI_{10/90}$ is < 5, there is increased concern for reproductive or developmental toxicity in humans, as there is limited separation in the doses causing adverse effects from those responsible for efficacy. If the $TI_{10/90}$ ratio falls between 5 and 20, the level of concern is unchanged. If the $TI_{10/90}$ ratio is > 20, there is decreased concern

[7] The TD_{10} (toxic dose or concentration) should be defined by the C_{max} (or other appropriate exposure metric) that produced the toxic reproductive or developmental response in 10% of a *responsive* or *sensitive* species, whereas the ED_{90} (efficacious dose or concentration) should be defined by the C_{max} (or other appropriate exposure metric) that produced the desired effect in 90% of the test species. These parameters can be estimated. Preferably, both the TD_{10} and ED_{90} would be defined in the same species. In some instances estimation of the ED_{90} can be based on in vitro cell inhibition studies (frequently seen for antibiotics and antineoplastic agents). Although less desirable, efficacy data can be derived from another species, but caution should be exercised in such situations. The same exposure metric should be used in the estimation of the TD_{10} and ED_{90} values. Scientific justification for the drug exposure metrics used for comparison should be provided.

because of the wide separation in doses causing adverse effects from those resulting in efficacy.

If there are data available to determine the $TI_{10/90}$ ratio in multiple species, assessment of the level of concern for this element should be based on an integrated analysis of data from all adequately studied species. The extent of concordance in the size of the $TI_{10/90}$ between species may increase, decrease, or leave unchanged the level of concern (i.e., the greater the concordance, the more likely concern will be increased). In the event of nonconcordance of the TI ratios between multiple test species, the nature of the positive signals observed and the relevance of the endpoint and test species to the human condition should be considered before making an assessment. In the event that one species is considered inappropriate to the analysis, the evaluation should be performed without reference to that species.

Biomarkers as a Benchmark

There may be circumstances in which an effect on a biomarker is consistently seen in multiple species at doses lower than the NOEL for demonstrable reproductive/developmental toxicity. If there is an effect on this biomarker at or below the therapeutic dose in humans, there is increased concern for reproductive or developmental toxicity in humans. If this biomarker is responsive to the drug in humans, can be monitored, and is not affected at the therapeutic dose, there may be decreased concern.

Similarity between Pharmacologic and Reproductive Developmental Toxicologic Mechanisms

If a positive signal is an extension of, progression of, or related response to the intended pharmacologic effect of the drug (e.g., delay of parturition by drugs known to suppress uterine smooth muscle contractility or hypotension in the offspring of dams treated during late gestation with a drug known to lower blood pressure), there is increased concern for reproductive or developmental toxicity in humans. There is less concern if the positive signal is attributed to an animal-specific pharmacological response, even though it may be an extension of the pharmacologic effect of the drug (e.g., pregnancy loss in rats due to hypo-prolactinemia).

d. Concordance between the Test Species and Humans

Concordance between the test species and humans should be evaluated with respect to: (a) the metabolic and drug distribution profiles, and (b) the general toxicity profiles, and (c) biomarker profiles.

Metabolic and Drug Distribution Profiles

Drug distribution, elimination, and biotransformation (pathways and metabolites) in the test species and in humans should be compared. Quantitative differences in metabolic/drug distribution profiles between the test species and humans are often seen, and may not have important implications and should not be overemphasized. Reproductive and developmental toxicities induced by compounds whose metabolic and distribution profiles are very similar in animals and humans increases concern for reproductive or developmental toxicity in humans. For compounds with highly dissimilar metabolic or tissue distribution profiles in animals and humans, there is less concern if the toxic effect seen in the test species can be attributed to a metabolite or tissue distribution profile not seen in humans. For any other scenario, concern is unchanged.

When there are significant differences in drug distribution or metabolic profiles between several species, yet each test species demonstrates a positive signal for a reproductive or developmental toxicity, the toxicity is assumed to be attributable to the parent drug or a common bio-transformed product and concern is increased.

General Toxicity Profiles

If the overall toxicity profile of a drug in one or more test species with a positive signal is similar to that in humans, there is increased concern for reproductive or developmental toxicity in humans. If the overall toxicity profiles are dissimilar, there may be decreased concern. When general toxicology data are available for more than one species, the determination of the level of concern (increased, decreased, or unchanged) should be based on an assessment of each test species' ability to indicate human adverse effects in response to the drug.

Biomarker Profiles

When biomarker profiles are available for comparison, an approach similar to that described in the assessment of General Toxicity Profiles (previous section) may be useful.

e. Relative Exposures

When considering the relative exposure comparisons discussed below, more emphasis should be placed on a parameter within this factor when there is a scientifically plausible link between the exposure metric (or biomarker) and the adverse reproductive (or developmental) effect.

Kinetic Comparison of Relative Exposure

Comparison of systemic drug exposure at the NOEL for the reproductive or toxicity class in the test species to that in humans at the maximum recommended dose is a critical determination. This comparison should be based on the most relevant metric (e.g., AUC, C_{max}, C_{min}, BSA [body surface area] adjusted dose). In general, there is increased concern for reproductive or developmental toxicity in humans for relative exposure ratios (animal:human) that are $\leq$10, decreased concern for exposure ratios $\geq$25, and no change in concern for ratios between 10 and 25. When applicable, the relative exposure ratio should consider both the parent compound and its metabolites. For example, it is appropriate to combine parent and metabolite when both are pharmacologically active and the activity relates to the reproductive or developmental toxicity.

Where there are exposure data for multiple test species, the NOEL exposure for each should be compared to human exposure at the maximum recommended dose. If the exposure ratios are low ($\leq$10 fold) in multiple species with a positive signal, there is increased concern. If the exposure ratios are high ($\geq$25 fold), there is decreased concern. In the event a significant difference in relative exposures is observed between multiple test species, the appropriateness of the metric (for example, AUC, C_{max}) being used to define the inter-species exposure comparisons should be re-assessed. If an alternative metric fails to reduce the disparity between species, the assessment of concern should be based on the lowest ratio (i.e., in the most sensitive species).

Relative interspecies exposure data should be evaluated in light of species differences in protein binding (free drug concentration), receptor affinity (if related to the positive signal) or site specific drug concentrations. In the absence of meaningful differences between the test species and humans in these parameters, the interspecies comparisons should be based on total drug exposure.

Biomarkers as a Measure of Relative Exposure

The purpose of this relative exposure metric is to compare the dose causing a reproductive toxic effect in the test species to the therapeutic dose in humans, normalized to the doses causing a

response common to both species. In practice, this is done by taking the NOEL for the adverse reproductive or developmental effect and dividing by the dose at which the biomarker response is seen in the test species. This is compared to the human therapeutic dose divided by the dose at which the biomarker response is seen in the human. The ratio calculated for animals is then divided by the ratio calculated for humans. When this ratio of relative biomarker exposure (animal:human) is $\leq$ 10, there is increased concern for human reproductive or developmental toxicity. When this ratio is $\geq$ 25, there is less concern. When this ratio falls between 10 and 25, the level of concern is unchanged.

Where there are data to compute relative biomarker exposure ratios for multiple species, the level of concern assessment should be based on an integrated analysis of data from all adequately studied species. Where there are nonconcordant biomarker ratios between multiple test species, the relevance of the biomarker as expressed in the various species should be considered before making an assessment. If there is no scientific rationale for the disparity between species, the biomarker, as a measure of exposure, will be of questionable utility.

f. Class Alerts

Consideration of a *class associated effect* should be based on prior human experience for a drug with related chemical structure (parent or metabolite) or related pharmacologic effect, and with known reproductive or developmental outcomes in humans. There is increased concern for reproductive or developmental toxicity in humans when the drug is from a class of compounds known to produce adverse effects in the same toxicity class in humans and animals. There is decreased concern only in circumstances in which a class of compounds, although demonstrating adverse effects in animals, has been previously shown to have no related adverse effects on human reproduction or development. In the absence of adequate human reproduction or developmental data for a class, the level of concern is unchanged.

g. Summary/Integration of Positive Findings

Notes on the use of the Integration Tool (see Figure C end of document):

The factors discussed below are derived from a limited sample of pharmaceuticals where the clinical outcomes are reasonably well defined. CDER believes that using specific factors and benchmark values to assess the potential to increase risk to humans for adverse

reproductive and developmental outcomes will result in a more unbiased and uniform evaluation. CDER also believes this approach will help identify specific areas of additional information about a pharmaceutical that would be useful in more fully defining risk and allow specific analysis of areas of disagreement that influence the risk evaluation.

1. Where there is a positive finding in nonclinical or general toxicology studies for one of the seven classes of reproductive or developmental toxicity, there is a potential for increased human risk. In evaluating the level of increased risk, positive findings from each of the seven classes of reproductive and developmental toxicity should be assessed separately. All relevant information should be considered.

2. In evaluating the level of concern for each of the six factors in the overall assessment, the analysis should reflect the weight of evidence taking into account the quality and type of data under consideration for each factor (i.e., should not be merely an arithmetic summation of the contributory elements for each factor). For each factor there should be a determination of increased (+1), decreased (-1), or no change (0) in the level of concern.

3. The values for the six factors should then be summed to arrive at one of the following overall conclusions for each class of reproductive or developmental toxicity: (1) the drug is predicted to increase risk, (2) the drug may increase risk, or (3) the drug does not appear to increase risk of that class of reproductive or developmental toxicity in humans. Where there is sufficient information about the drug to assess each of the six factors within Figure C, a net value of $\geq +3$ suggests a drug is predicted to increase risk for that class of toxicity in humans, a value between +2 and –2 suggests that the drug may increase the risk, and a value ≤ -3 suggests the drug does not appear to increase the risk.

The summary risk conclusions for the outcomes of analyses using Figure "C" are:

Does Not Appear to Increase Risk: The drug is not anticipated to increase the incidence of adverse reproductive (or developmental) effects above the background incidence discussed in humans when used in accordance with dosing information in the product label.

May Increase Risk: The drug may increase the incidence of adverse reproductive (or developmental) events above the background incidence in

humans when used in accordance with the dosing information in the product label.

Predicted to Increase Risk: The drug is expected to increase the incidence of adverse reproductive (or developmental) events above the background incidence in humans when used in accordance with the dosing information in the product label.

GLOSSARY

ADME – absorption, distribution, metabolism, and elimination

Biomarker – a clinical or laboratory parameter that is known, or thought, to correlate with a toxicity outcome or with exposure

Class Alert – an adverse reproductive or developmental effect previously demonstrated in humans by closely related chemical entities or compounds with similar pharmacodynamic effects

Contributory Elements – specific items of information that contribute to the overall evaluation and conclusion for each factor of Figure C

Developmental Toxicity – any adverse effect induced prior to attainment of adult life. It includes effects induced or manifested in the embryonic or fetal period and those induced or manifested postnatally. These are generally adverse effects that affect the F_1 generation and are divided into four endpoints, mortality, dysmorphogenesis, alterations to growth, and functional toxicities.

Factor – for purposes of this guidance, a factor is one of the six components used to evaluate the level of concern for a positive signal in a class of developmental or reproductive toxicity to determine whether there is an increase (assigned value of +1), a decrease (-1), or no change (0) in the overall concern for that class of toxicity. There are six factors: (1) signal strength, part I; (2) signal strength, part 2; (3) pharmacodynamics; (4) concordance between the test species and humans; (5) relative exposures; and (6) class alerts. They are all portions of the Integration Tool (Figure C) and are discussed in Section 4.

Fertility – reproductive competence

Lactation – the secretion of milk or the period of milk secretion

Malformation – a permanent alteration (anomaly) in which there is a morphologic defect of an organ or a larger region of the body, resulting from an abnormal developmental process. They generally occur at a low frequency in the control population and/or will adversely affect survival, growth, or development of functional competence.

Maternal (Paternal) Toxicity – toxicity to the mother (maternal) or the father (paternal) in a reproductive toxicology study, but not necessarily a toxicity to reproductive function

Parturition – labor and delivery

Positive Signal – a treatment related reproductive or developmental toxicity

Rare Event – an endpoint that occurs in less than 1 percent of the control animals in a study and in historical control animals

Reproductive Toxicity – structural and/or functional alterations that may affect reproductive competence of the F_0 generation. These are divided into three classes—fertility, parturition, and lactation.

Therapeutic Index – for the purpose of this document, the ratio of the dose that induces a toxicologic effect in approximately 10 percent of the treated animals (TD_{10}) compared to the dose that brings about the intended result of the therapeutic in 90 percent of the treated animals (ED_{90})

Toxicologic Effect – any adverse effect of a therapeutic

Variation - an alteration that may occur at a relatively high frequency and/or represents a retardation in development, a transitory alteration, or a permanent alteration not believed to adversely affect survival, growth, development, or functional competence.

APPENDIX A

SAMPLE SCENARIOS AND RISK CONCLUSIONS FOR SITUATIONS IN WHICH THERE ARE NO POSITIVE FINDINGS FOR A CLASS OF REPRODUCTIVE OR DEVELOPMENTAL TOXICITY (ASSESSMENTS USING FIGURE B)

Case 1. The animal species and dose selections were considered appropriate, there is no class alert for the drug, and no positive signals were observed for any class of developmental toxicity.

Summary Risk Conclusion 1. Based on studies in animals, there does not appear to be an increased risk for adverse developmental effects in humans.

Case 2. No positive signals were observed for any class of reproductive toxicity and there are no human data for the drug. However, other drugs in the same pharmacologic class have demonstrated adverse reproductive effects in humans (i.e., a class alert).

Summary Risk Conclusion 2. The risk for adverse reproductive effects in humans is unknown. Although no effects were observed in adequately conducted reproductive toxicity studies in animals, and there is no information about adverse reproductive effects of the drug in humans, adverse reproductive effects have been observed in humans with related drugs. (should specify the type of adverse effects observed in humans with other members of the class and the basis for the class designation—e.g., chemically or pharmacologically related).

Case 3. The available developmental toxicity studies are considered to lack predictive value because exposures to the drug in animal studies conducted at the MTD were not considered adequate when compared to the maximum exposure in humans.

Summary Risk Conclusion 3. The risk for adverse developmental effects in humans is unknown. Although there were no observed adverse developmental effects in adequately conducted toxicity studies in animals, exposures achieved in the animal studies may not have been adequate to fully evaluate the potential for the drug to increase the risk of reproductive or developmental toxicity in humans.

Case 4. The animal models were not considered adequate to test the drug because the test species in the reproductive toxicity studies lacked the cellular receptor responsible for the pharmacologic activity of the drug in humans, or did not demonstrate a toxicity or metabolite profile similar to that in humans.

Summary Risk Conclusion 4. The risk for adverse reproductive effects in humans is unknown. Although there were no observed adverse effects in animal reproductive toxicity studies, there remains some concern for increased risk of adverse reproductive effects in humans exposed to the drug because the test species may not be predictive of the human condition.

Case 5. In animal studies considered appropriate for predicting the human response, and at exposure levels significantly greater than expected in humans, there was a positive signal in one class of developmental toxicity and no observed adverse effects in a related class of developmental toxicity. The positive signal was evaluated using Figure C and it was concluded that the drug may increase the risk of that class of toxicity in humans.

Summary Risk Conclusion 5. Based on studies in animals, the drug may increase the risk of [the class of developmental toxicity in which toxicity was observed] in humans. Although no findings were observed for [type of developmental toxicity in same category] there may be some relationship between the incidence of [the class of toxicity in which toxicity was observed] and [the related class of toxicity not observed].

FIGURE A. OVERALL DECISION TREE FOR EVALUATION OF REPRODUCTIVE/DEVELOPMENTAL TOXICITIES

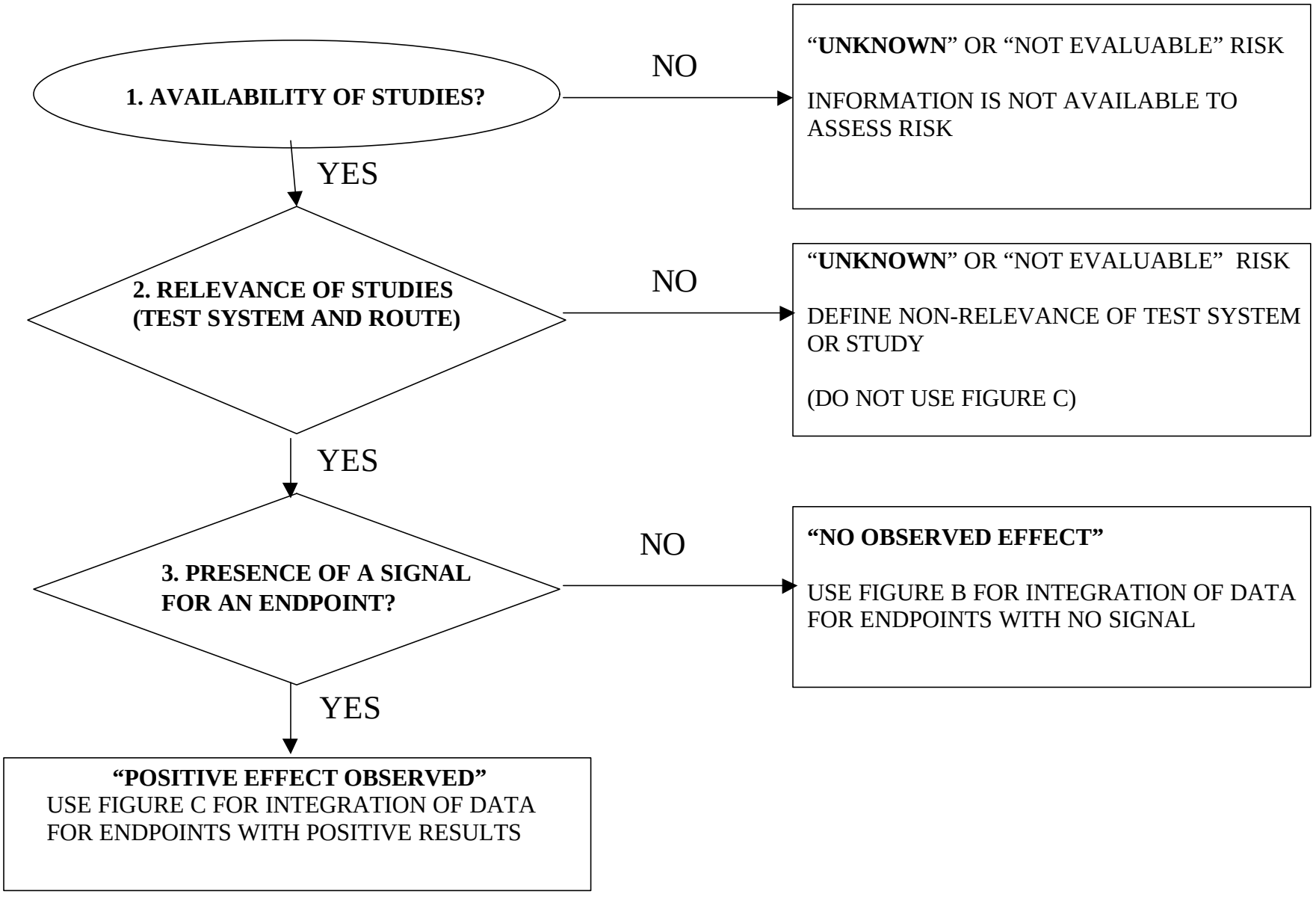

FIGURE B. DECISION TREE FOR REPRODUCTIVE/DEVELOPMENTAL TOXICITIES WITH NO SIGNAL

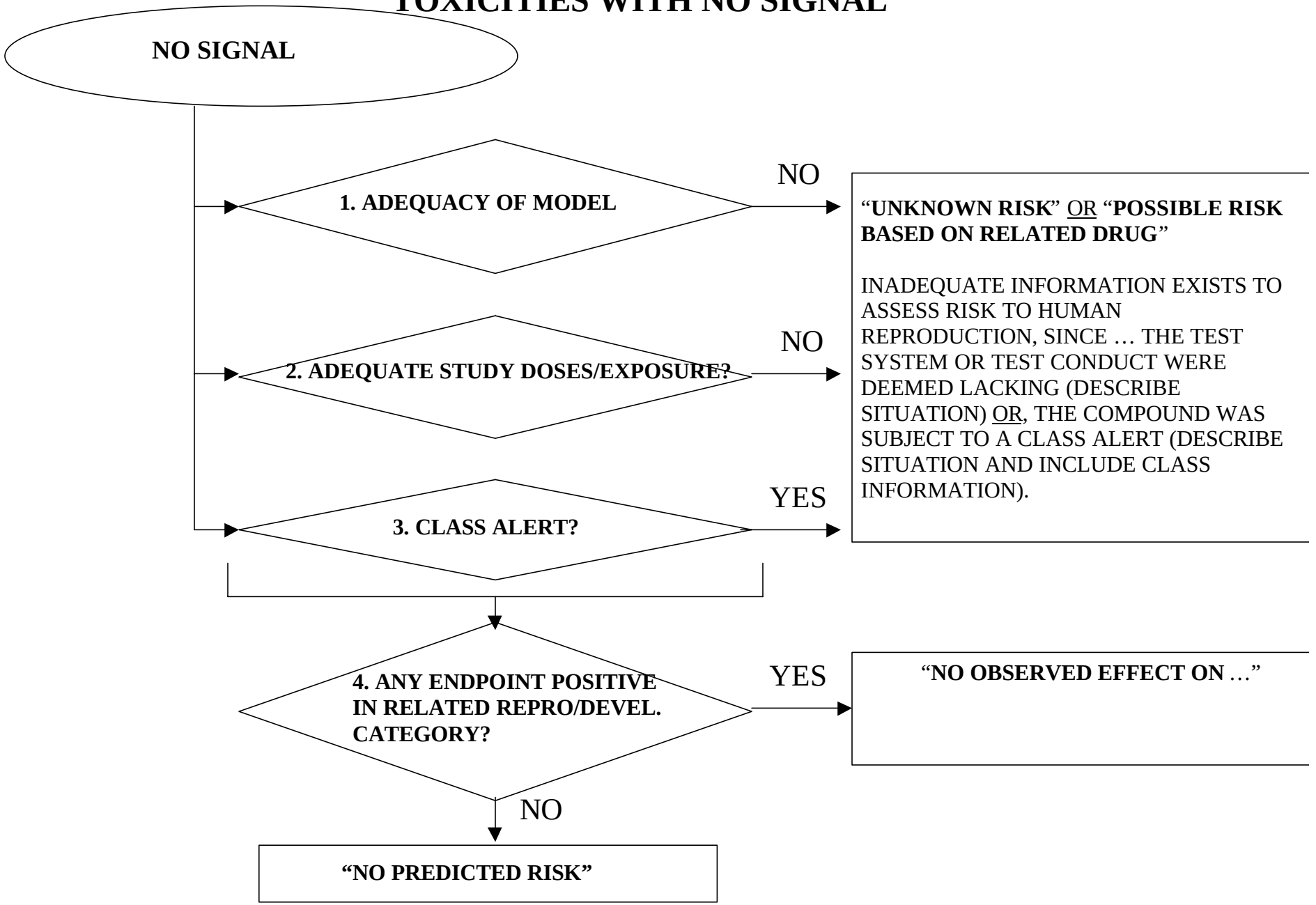

FIGURE C - INTEGRATION TOOL FOR REPRODUCTIVE OR DEVELOPMENTAL TOXICITIES WITH A POSITIVE SIGNAL