Human relevance of developmental animal toxicity data of pharmaceuticals from the perspective of the European Teratology Society

7th Workshop on the Terminology in Developmental Toxicology
May 04 to 06th, 2011
Berlin

S. Barbellion
"any drug administered at the proper dosage, and at the proper stage of development to embryos of the proper species-and these include both vertebrates and invertebrates-will be effective in causing disturbances in embryonic development“

Karnofsky, 1964
Introduction

- Concordance (=agreement)
  - Thalidomide tragedy
  - Regulatory testing
    - ICH guidelines for pharmaceuticals
    - 2 species
  - General principles of Teratology (Wilson)

- Relevance (=pertinence)
  - Toxicology testing
  - Risk assessment
  - Regulatory aspects
    - Integrative Assessment

- Significance (=importance, meaning)
  - Terminology and classification
  - Teratology findings
Four manifestations of developmental toxicity

- Structural malformations
  - 3 to 6% of birth defects (Nelson, Holmes 1989)
    - 28% genetic causes
    - 23% multifactorial inheritance
    - 3% uterine factors
    - 3% toxicants
    - 43% unknown

- Deaths
  - (miscarriage)

- Growth retardation
  - Low birth weights (5% of babies); predictor of susceptibility to certain chronic disease

- Functional deficits
  - Mental retardation (lead, alcohol)
Principles of Developmental Toxicology

- Incidence and severity dependent on dose (and route of administration)
  - Dose-related effect (threshold)
    - Inorganic arsenic (route of administration)
  - More than one manifestations can occur (or one can hide the other)

- Mechanism of action
  - Safer drugs can be designed, safer use
  - Predictive toxicology

- Genetic background and interaction with environment
  - Mother and embryo
  - Difference is metabolism or pharmacokinetics
    - alcohol

- Stage of development at the time of exposure
  - Critical period of development (Wilson, 1973)
    - Thalidomide limb defects between Days 24 to 34 post-fertilization
Concordance

- **Animal-to-human**
  - Human experience through epidemiology in needed
  - 70-80% concordance with either rodents or rabbits
    - Responses between animals and humans can be different; but evidence of developmental toxicity that would have elicited regulatory action
  - Rodent studies are the most concordant, but also the most non concordant responses (Shardein, 1985)
    - Mycophenolic acid
    - Oral Isotretinoin
  - Olson et al. (1998, 2000)
    - Animal Toxicity vs. Human toxicity (HT) of pharmaceuticals during clinical trials
      - 71% positive HT concordance with rodent and non rodent species

- **Human teratogens found positive in at least one animal species**
  - 40-50 environmental factors (agents and pathogens, about 25 drugs)
  - Discrepancies between number of chemicals that have DevTox and number of known human developmental toxicants
    - Regulatory system
    - Shepard’s catalog, reprotox database

### Table 1: Comparison of developmental toxicity detection in rodents and rabbits with human response

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Rodent</th>
<th>Rabbit</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclophosphamide</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Diazepam</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Diethylstilbestrol</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Phenytointrimethadione</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ethanol</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Lithium</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>Methymercury</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>13-cis-Retinoic acid</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Testosterone</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Thalidomide</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Valproic acid</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Warfarin</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Fumonisins B1</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Methimazole</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Busulfan</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Enalapril/captopril</td>
<td>+/-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Polychlorinated triphenyls</td>
<td>+/-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Cocaine</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Misoprostol</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Penicillamine</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Toluene</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+ indicates developmental toxicity in that species; – indicates lack of response; +/- indicates an equivocal response, or a response that might not have been interpreted as indicative of a specific response. Note that a + rating does not necessarily mean that the same response was elicited in all species, but that some significant unequivocal manifestation of developmental toxicity was observed.

Daston et al., 2010

---

**Sanofi Aventis**

L’essentiel c’est la santé.
### Table 4. Predictability of animal models to concordant human malformations.

<table>
<thead>
<tr>
<th>Teratogen</th>
<th>Reference malformation</th>
<th>Concordant</th>
<th>Nonecordant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol</td>
<td>Craniofacial, limb, CV</td>
<td>Mouse, dog</td>
<td>Rat, guinea pig, pig</td>
</tr>
<tr>
<td>Androgenic/progestogenic hormones</td>
<td>Pseudohermaphroditism (♀)</td>
<td>Mouse, rat, guinea pig, hamster, rabbit, dog, pig, primate</td>
<td></td>
</tr>
<tr>
<td>Anticancer antimetabolites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aminopterin</td>
<td>Skeletal</td>
<td>Rat</td>
<td>Dog, pig</td>
</tr>
<tr>
<td>Fluorouracil</td>
<td>Multiple visceral</td>
<td>Mouse, rat, guinea pig</td>
<td>Rabbit, primate</td>
</tr>
<tr>
<td>Methotrexate</td>
<td>Skeletal</td>
<td>Rabbit, cat</td>
<td>Mouse, rat, primate</td>
</tr>
<tr>
<td>Cytarabine</td>
<td>Limb, ear</td>
<td>Rat</td>
<td>Mouse</td>
</tr>
<tr>
<td>Anticancer alkylating agents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Busulfan</td>
<td>Multiple visceral</td>
<td></td>
<td>Mouse</td>
</tr>
<tr>
<td>Chlorambucil</td>
<td>Urogenital</td>
<td></td>
<td>Mouse, rat</td>
</tr>
<tr>
<td>Cyclophosphamide</td>
<td>Digits</td>
<td>Mouse, rat</td>
<td>Rabbit, primate</td>
</tr>
<tr>
<td>Mechlorethamine</td>
<td>Renal, limb, ear</td>
<td>Rat, rabbit, ferret</td>
<td>Mouse</td>
</tr>
<tr>
<td>Anticonvulsants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydantoin</td>
<td>Facial, mental</td>
<td>Mouse</td>
<td>Rat, rabbit, primate</td>
</tr>
<tr>
<td>Diones</td>
<td>Facial, mental</td>
<td></td>
<td>Mouse, rat</td>
</tr>
<tr>
<td>Valproate</td>
<td>CNS</td>
<td></td>
<td>Mouse, rat</td>
</tr>
<tr>
<td>Antithyroid agents</td>
<td>Hypothyroidism</td>
<td>Mouse, rat, guinea pig, rabbit</td>
<td></td>
</tr>
<tr>
<td>DES</td>
<td>Uterine lesions</td>
<td>Mouse, rat, primate, ferret</td>
<td>Hamster</td>
</tr>
<tr>
<td>Methylmercury</td>
<td>Microcephaly, mental</td>
<td>Mouse, rat, cat</td>
<td></td>
</tr>
<tr>
<td>Thalidomide</td>
<td>Limb</td>
<td>Rabbit, primate</td>
<td>Mouse, rat, hamster, dog, cat, pig, ferret</td>
</tr>
<tr>
<td>Lithium</td>
<td>CV</td>
<td></td>
<td>Mouse</td>
</tr>
<tr>
<td>o-Penicillamine</td>
<td>Skin lesion</td>
<td>Rat</td>
<td>Hamster</td>
</tr>
<tr>
<td>Streptomycin antibiotics</td>
<td>Inner ear</td>
<td></td>
<td>Rat</td>
</tr>
<tr>
<td>Vitamin A analogs</td>
<td>CV, ear, brain</td>
<td>Rat, mouse, hamster, dog, primate</td>
<td>Rabbit, guinea pig, pig</td>
</tr>
</tbody>
</table>
Concordance

Concordance animal-to-animal

- Hurt et al. (2003) - teratogenicity
  - 61% of (a series of) Veterinary drugs showed teratogenicity in any one of the species.
  - 100% in rats and rabbits together

- Janer et al. (2008) – developmental toxicity
  - Assessed the added value of a second species (rabbit) when rat data are available
  - ICH S9 (oncologic compounds), S6 (biologics)
  - Overall same sensitivity across species with regards to developmental toxicity
  - Direct vs. indirect maternally-mediated effects was problematic in the interpretation
Predictivity of developmental toxicity

- Nonanimal models
  - Rodent embryos in culture
  - Mouse embryonic stem cells and hES
  - Free-living embryos (xenopus, zebra fish)
  - Primary cultures of embryonic tissues

- Teratogenicity Screening Strategy
  - Prerequisites, high throughput
  - Level of concern
  - Warning for subsequent testing (in vivo screening)
Animal testing is considered to be relevant for predicting human toxicity

- Assume humans are more sensitive
- NOEL values can be used to predict safe levels in humans
  - Concern when effects within 20-fold the therapeutic blood level?
  - 100x margin is sufficient?
- Mechanistic study to demonstrate that a finding in DART studies is not relevant to humans
- Hierarchization (and extrapolation) of animal teratology findings?

---

**Suggested recommendation to clinicians based on experimental data**

<table>
<thead>
<tr>
<th>Animal exposure level</th>
<th>Compound-related animal malformations</th>
<th>Compound-related animal embryofetal toxicity (excluding malformations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High human risk expected</td>
<td>Human risk cannot be excluded</td>
<td></td>
</tr>
<tr>
<td>Animal exposure level several-fold higher than human therapeutic dose exposure</td>
<td>Human risk cannot be excluded</td>
<td>No human risk expected</td>
</tr>
</tbody>
</table>

Guittin et al., 2000
Relevance

- Importance of rigorous and relevant testing
  - Quality of the data, relevance of findings
    - Good dose selection for definitive DART studies (avoid marked maternal toxicity)
    - Adequate study design and number of animals
  - Differences in sensitivity between species required attention
    - Relevance of the effects (mechanistic study)
    - Metabolism differences, pharmacokinetics
  - Placental differences (Carney et al. 2004)
    - Inverted yolk sac (histiotrophic nutrition, trypan blue in rats)
Relevance

- Importance of rigorous and relevant testing
  - Maternally-mediated toxic effects
    - Finding that occur in the absence of MatTox are most relevant to the clinical situation where MatTox is unlikely at therapeutic doses (FDA)
    - Sponsors should be able to support claims that DevTox is due to MatTox (EMA)
      - Presence of maternal and developmental toxicity does not change the level of concern for human risk
      - Demonstrate that the developmental toxicity is not relevant to humans not simply that is due to maternal toxicity
  - Assess human relevance of a maternally mediated MOA
    - Diflunisal-induced matenal anemia (Clark, 1984)
    - Potent hERG channel blockers produced fetal death in conventional studies, stage-specific malformations in single administration (Danielsson, ILSI/HESI WS 2010)
  - Weight-of-evidence evaluation of data
    - Plausibility if a causal link e.g. decrease maternal food consumption and BW gain: decrease fetal BW (plausible); fetal malformations (not plausible)
    - Temporal correspondence
    - Individual animal correspondence (association of maternal and litter effects)
    - Potential interactions
    - MOA

- Historical control data
  - Genetic background and drift

Decision logic for assessing the human relevance of a maternally mediated mode of action (MOA) (Carney, 2010)
Relevance and significance of findings

- Primary concern has been the terminology used to describe structural changes in the offspring
  - Misclassification or inconsistencies in the use of terms
  - Comparative atlas of malformation, images database
    - Guittin et al.
    - DevTox website
  - Classification
    - Malformation and Variation, Chahoud et al., 1999
    - Berlin WS
  - Terminology
    - Wise et al., 1997
    - Makris et al., 2009
  - Grading, severity and adversity
    - Paumgartten et al., 2009
- Integrative Assessment of nonclinical and clinical findings
  - EMA guideline, 2008
  - FDA guidance, 2001
EMA Integrative Assessment

- **Nonclinical Assessment**
  - Reproductive Tox studies and all pharmacological and toxicological studies
  - Choice of species (2 species for EFD, at least one responsive to the PD effect)
  - Pharmokinetics (relevance of the species)
    - Placental transfer and Milk excretion study is of value for the assessment
  - Route of administration
  - Toxicokinetics in pregnant animals
    - Metabolite
    - Comparison of toxic and PD effective dosage, animal to human exposure ratio
  - Dose Levels
    - Minimal maternal toxicity (magnitude and nature to be considered for relevance)
    - NOAEL
  - Mechanism (desirable when reproductive toxicity identified)
  - Class alert

- **Evaluation process**
  - Identified lack of data
  - No effect detected
  - Effects detected
    - Recognition of an effect
      - Incidence, rarity, dose-response relationship
    - Cross-species concordance
      - Increase the concern
    - Type of effects
    - Multiplicity of effects

- **Clinical Assessment**
## EMA Integrative Assessment - Pregnancy

<table>
<thead>
<tr>
<th>Non clinical data</th>
<th>Effects detected</th>
<th>No effects detected</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Human data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demonstrated human teratogenicity (or fetotoxicity)</td>
<td>Proven risk in humans <em>Labelling</em> [1] <em>See also decision scheme on Contraindication</em></td>
<td>Proven risk in humans <em>Labelling</em> [1] <em>See also decision scheme on Contraindication</em></td>
</tr>
<tr>
<td>Supposed or suspected human teratogenicity (or fetotoxicity)</td>
<td>Strong suspicion of risk in humans <em>Labelling</em> [2]</td>
<td>Risk is possible in humans <em>Labelling</em> [3]</td>
</tr>
<tr>
<td>None or less than 300 prospective exposed pregnancies with known outcome in the 1st trimester and no increased rate of malformation identified</td>
<td>Risk is possible in humans, not confirmed <em>Labelling</em> [4]</td>
<td>Malformative risk unlikely in humans, but low evidence <em>Labelling</em> [5]</td>
</tr>
<tr>
<td>At least 300 prospective exposed pregnancies with known outcome in the 1st trimester and no increased rate of malformation identified</td>
<td>Malformative risk unlikely in humans, but low evidence <em>Labelling</em>[6]</td>
<td>Malformative risk unlikely in humans with moderate to substantial evidence <em>Labelling</em>[7]</td>
</tr>
<tr>
<td>At least 1000 prospective exposed pregnancies with known outcome in the 1st trimester and no increased rate of malformation identified</td>
<td>Malformative risk unlikely in humans with strong evidence <em>Labelling</em>[8]</td>
<td>Malformative risk unlikely in humans with strong evidence <em>Labelling</em>[8]</td>
</tr>
</tbody>
</table>

* Insufficient data are considered as effects detected
Contraindication in pregnancy

Documentation of studies to be provided by the innovator company, as well as literature data

Sufficient Human Experience?

Yes

Evidence of Risk?

Yes

Treatment Avoidable? Postponable?

Yes

Contraindication in Pregnancy in 4.3 and 4.6

No

No

Information in 4.6 and 5.3

No

Yes

Relevant Risk from Non-Clinical Studies

Information in 4.6

Information in 4.6 Stringent wording Case-by-Case (also 5.3)
Integrative Assessment tool

- Integrative Assessment (IA) for Evaluating Data for Potential Human Developmental and Reproductive Toxicology (DART) Risk
- A process for evaluating potential human developmental and reproductive toxicology (DART) risk of a compound using available nonclinical and clinical data.
  - Based on procedures proposed by US FDA and EMEA
  - Uses a weight-of-evidence approach
IA: Data Categorization

- Reproductive Endpoints (F0)
  - Fertility and Fecundity
  - Parturition
  - Lactation

- Developmental Endpoints (F1)
  - Mortality
  - Dysmorphogenesis
  - Alterations in growth
  - Functional toxicity

- Additional Data Sources
  - Class alerts
  - Other in-house compounds
  - Human and animal exposure data
  - Published literature
  - Position papers
  - Additional nonclinical safety studies
  - Etc.
Components of the IA

- Populate templates for each endpoint assessed as appropriate:
  - Template A: Data adequacy
  - Template B: Decision process for endpoints with no adverse effects
  - Template C: Decision process for endpoints with DART effects
Flowchart A: Data Adequacy

1. Necessary studies conducted?
   - NO: State that no information is available for assessing risk. If warranted, propose additional studies.
   - YES:
     2. Test system and route relevant to humans?
        - NO: Describe situation as to relevance of test system and route. Do not use Flowchart C.
        - YES:
          3. Adverse DART effect for an endpoint?
             - NO: Use Flowchart B for endpoints with no adverse DART effects.
             - YES: Use Flowchart C for endpoints with adverse DART effects.
Flowchart B: For Endpoints with No DART Effects

1. Models predictive?
   - NO
     - No adverse DART effect
   - YES
     - 2. Studies adequate?
       - NO
         - Inadequate information to fully assess risk to humans because-- (Describe situation and stop assessment for specific endpoint)
       - YES
         - 3. Class alert?
           - NO
             - State endpoint affected
           - YES
             - 4. Any related endpoint positive?
               - NO
                 - No human Class effect
               - YES
                 - State endpoint affected
               - Predicted Risk

Evaluate Class information (human)

Human Class effect
Flowchart C*: Integration of Available Data

Endpoints
A. Reproductive Toxicity
1. Fertility & fecundity
2. Parturition
3. Lactation

B. Developmental Toxicity
1. Developmental mortality
2. Dysmorphogenesis
3. Alterations to growth
4. Functional toxicity

Endpoint with Adverse Effect

SIGNAL STRENGTHS
1. PD
2. ADME
3. EXPOSURE
4. CLASS, ALERTS

CONCERN

SIGNIFICANT CONCERN

NO KNOWN CONCERN

+4 to +7

-3 to +3

-4 to -7

DATA INTEGRATION PROCESS

HUMAN DATA
ANIMAL DATA

*Adapted from FDA, 2001

sanofi aventis
L’essentiel c’est la santé.
Template C: Adverse Effects Overview

- Assess each endpoint for which an adverse effect was observed according to 7 factors.
  - Signal Strength 1 (Multiplicity of effects, Adverse effects as a function of time)
  - Signal Strength 2 (Cross-species concordance, Parental toxicity)
  - Signal Strength 3 (Dose-response relationships, Rare events)
  - Pharmacodynamics
  - Species-Human Concordance
  - Animal:Human Relevant Exposure
  - Class Alerts

- Definitive studies carry more weight-of-evidence than range-finding studies. However, if the definitive study is conducted at lower doses than the range-finder, the range-finding study may still carry weight and may be included in the assessment.
Determine level of concern for each of the 7 factors and assign score:

- +1 for increased concern
- 0 for no change in concern
- -1 for decreased concern

Note:

- It is important to use a weight-of-evidence approach when assessing the level of concern for a particular factor, since some of the factors contain multiple points of consideration.
- After all 7 factors for a given endpoint have been determined, add the scores from all factors together to arrive at an overall risk value for the endpoint in question.
Template C: Adverse Effects Overview (cont.)

- Sum factors scores for each endpoint to determine concern for humans:
  - +4 to +7 = Significant concern for human risk
  - -3 to +3 = Concern for human risk
  - -4 to -7 = No known concern for human risk
Assessment relies mainly on nonclinical data until submission
  - Continuum and interactive process
  - Integrative analysis based on (full) evaluation of study reports
  - Fetal findings not reviewed

Clinicians use the data as indicators of potential outcomes in patients, while the scientist groups use the data as signal indicators of disrupted development.

Clinicians are not aware of the principles of the conduct of regulatory studies (dose selection, or use of effect doses in risk assessment estimate).

RMP (risk management plan)
  - Rare events should be mentioned