Reproductive Toxicology 27 (2009) 8–13

Contents lists available at ScienceDirect

Reproductive Toxicology

journal homepage: www.elsevier.com/locate/reprotox

Harmonization of terminology in developmental toxicology: The quest for a more precise description and a harmonized classification of fetal observations☆

ARTICLE INFO

Keywords:
Classification
Grey zone anomalies
Malformation
Variation
Developmental toxicology
Reproductive toxicology
Harmonization
Terminology

ABSTRACT

Harmonization of terminology in developmental toxicology is a prerequisite to ensure a better risk assessment of chemicals. As part of an international effort of the International Programme on Chemical Safety (IPCS) to harmonize terminology in developmental toxicology, workshops have taken place in Berlin since 1995. This publication reports the main outcomes of the Fifth and Sixth Berlin Workshops held in 2005 and 2007, respectively. The objective of the Fifth workshop was to discuss a draft international proposal for updating the glossary of descriptive terms for fetal abnormalities put forward by Wise et al. [Wise LD, et al. Terminology of developmental abnormalities in common laboratory mammals (version 1). Teratology 1997;55:249–92]. The participants were asked to classify the new external, visceral and skeletal observations included within this new version 2 of Terminology of Developmental Abnormalities in common Laboratory Mammals according to the two-category scheme (malformation and variation) agreed at previous Berlin workshops. The discussions held during the Sixth Workshop were mainly focused on the causes of uncertainty and low agreement regarding classification of some fetal observations as malformations or variations. Lack of precision in descriptive terms and insufficient knowledge of the postnatal consequences of fetal observations had been identified as major causes of uncertainty and lower agreement among evaluators regarding the classification of "grey zone anomalies", i.e. abnormalities that do not fit readily into one of the two categories (malformation or variation). Imprecise anatomical terms, observation terms that are too broad, lack of information on severity and the use of different terms for the same change or different severities of the same change, were found to be the main reasons that descriptive terms are often not sufficiently precise to allow accurate classification of findings. It was agreed that provision of additional information, including sub-location within the affected structure, more detailed description of the nature of the change, in conjunction with presentation of photographs wherever possible, and a grading for severity would make descriptive terms more precise, thereby reducing misclassifications. A better knowledge of the adversity and postnatal consequences of fetal observations was considered as the key issue for achieving a substantial reduction in the number of misclassifications and grey zone anomalies. The urgent need for additional research along this line as a prerequisite for a better risk assessment was emphasized by the participants.

1. Background

In the first Berlin Workshop, held in 1995, the International Federation of Teratology Societies (IFTS) terminology of developmental abnormalities in laboratory animals was discussed, and a first version of the IFTS glossary of descriptive terms was published [1]. The focus of the second Workshop, held in 1998, was the classification of fetal observations. After weighing the advantages and disadvantages of classification, it was agreed that classification of observations can be useful for risk assessment. A classification system based on only two categories, malformations and variations, was then advanced [2]. During the third Workshop, held in 2000, the experts discussed the results of a survey on the classification of skeletal anomalies as malformations or variations that had been sent to developmental toxicologists [3]. The term "grey zone anomalies" was used for those abnormalities that did not fit readily into one of the two categories (malformation or variation) [3]. In the fourth workshop, held in 2002, the discussion was focused on the results of a similar survey on classification of external and soft tissue anomalies [4]. In both surveys, a high agreement was reached among evaluators regarding classification of most anomalies described by IFTS glossary terms.

The discussion during the third and fourth workshops was focused on terms for which there was disagreement or uncertainty regarding classification. Fetal observations described by these terms were thereafter referred to as "grey zone anomalies". Among the possible reasons for lower agreement among evaluators on classification of terms as malformation or variation, the attendees of both workshops identified imprecise descriptive terms, insufficient knowledge of the postnatal consequences, theoretical terms that are unlikely to occur in isolation, and the possibility of observing a range of severity as key factors.

☆ Main outcomes of the 5th and 6th Workshops on the Terminology in Developmental Toxicology, Berlin, 2005 and 2007.

0890-6238/$ – see front matter
doi:10.1016/j.reprotox.2008.12.001

2. Outcome of the Fifth Berlin Workshop on the Terminology in Developmental Toxicology, Berlin, 27–29 October 2005

During the 5th Berlin workshop, discussions were concentrated on a draft proposal for updating the first version of the IFTS International Glossary ("Terminology of Developmental Abnormalities in Common Laboratory Mammals") published in 1997 [1]. It was recommended that contributions by three international teratology societies (US, Japan and Europe) should be harmonized to form the basis for a new version of the glossary.

Participants in the workshop were asked to classify the new external, soft tissue and skeletal observations, included within the European proposal, for a new Version of the Terminology of Developmental Abnormalities in Common Laboratory Animals, according to the scheme agreed at previous Berlin Workshops. The options for classification were "M" (Malformation), "V" (Variation), "U" (cannot decide between Malformation or Variation) and decisions were based on the definitions of malformation ("a permanent structural change that is likely to adversely affect the survival or health of the species under investigation") and variation ("a change that occurs within the population under investigation and is unlikely to adversely affect survival or health.") previously agreed at the Second Berlin Workshop [2]. Results of this survey were analysed using a formula to give an index of agreement (IA), where $IA = [(M - V)/(M + V + U) \times 100]$. The maximum scores for agreement were +100 for malformations and −100 for variations. The minimum score for agreement (i.e. the maximum score for disagreement) was 0. Good agreement was indicated when IA was between 100 and 75. Medium agreement was indicated when IA was between 25 and 75 and poor agreement when IA was between 0 and 25.

The new terms in this updated version of the glossary and the outcome of this survey will be published in the DevTox web site (www.devtox.org).

3. Outcome of the Sixth Berlin Workshop on the Terminology in Developmental Toxicology, Berlin, 25–27 October 2007

During the 3rd and 4th Berlin workshops [3,4], lack of precision in descriptive terms and insufficient knowledge of the postnatal consequences of fetal observations had been identified as major causes for uncertainty and lower agreement among evaluators regarding the classification of "grey zone anomalies". The discussions held at the 6th workshop were focused mainly on these two topics.

3.1. Lack of precision of descriptive terms

Descriptive terms for fetal anomalies are in most cases two-term statements, with one term (*anatomical term*) that identifies the altered anatomical structure and a second term (*observation term*) that indicates how the structure is altered. Examples would be "fibula bent", "hyoid body absent", and "zygomatic misshapen", where "fibula", "hyoid body" and "zygomatic" are the anatomical terms and "bent", "absent" and "misshapen" are the observation terms. The main reasons why descriptive terms are often not sufficiently precise to allow accurate and harmonized classification of fetal observations are imprecise anatomical terms, use of observation terms that are too broad, lack of information on severity and the use of different terms to describe the same change or different severities of the same change.

3.1.1. Imprecise anatomical terms

Imprecise identification of the structure or the part of the structure that is altered may be misleading. Examples of this would be using the descriptive term "maxilla misshapen" instead of "zygomatic process of maxilla misshapen", when only the *zygomatic process* of the bone is altered, or using "humerus misshapen" when only the "deltoid process of the humerus" is altered (Fig. 1). In both cases, the anatomical term is restricted to the main bone name, when only one small part of the bone is in fact affected.

3.1.2. Use of observation terms that are too broad

An imprecise description of the fetal anomaly also occurs when the observation term is too broad and can encompass a whole range of different types of change. The observation term "misshapen", for instance, can be used to describe virtually all alterations from the normal shape, i.e. any "variation" or "malformation". A more precise description of the shape change is often needed for classification. Therefore, "misshapen" should be further qualified ("modified") by descriptors such as "bulbous", "pointed caudally", "rounded", "with dorsal protuberance", wherever possible. Similarly, more information is necessary to aid decision making regarding classification of "discoloured" or "malpositioned" structures. In these cases, knowledge of the extent and nature of the colour change and description of the actual location of the structure would be helpful.

3.1.3. Lack of information on severity

Lack of information on the severity of the observation may also preclude an accurate classification of reported findings. Observation terms such as "small", "short", "long", "thin", "thick", "narrow", "large", "dilated" and "bent" cover a wide range of possible appearances, and structures described by these terms can range in appearance from something that is only marginally outside the "normal range" to something that is severely malformed (Fig. 2).

3.1.4. Use of different terms to describe the same change or different severities of the same change

Lack of precision in descriptive terms may also arise due to use of different terms to describe the same change or different severities of the same change. For example, depending on the severity, a skull closure effect in rat fetuses can be described as meningocele, meningoencephalocele or exencephaly. In this sequence, use of a grading scale for skull closure defects, instead of recording and analysing each one of these mechanistically related changes under a separate descriptive term, seems to be a more powerful approach for evaluating dose–response relationships [5]. In a similar way, the possibility that cleft palate, misshapen palatine rugae and wider spacing of the two parts of the palatine bone are different severities of the same underlying change was demonstrated in mice treated with triphenyltin hydroxide (TPTH) [6]. A study of the effects of teratogenic agents on the development of the rat fetus skeleton also suggested that those changes that are separately described as "vertebral ossification centre partially split", "vertebral ossification centre dumbbell shaped", "vertebral ossification centre bipartite" and "vertebral ossification centre hemicentric", are all manifestations of the same underlying mechanism and represent different degrees of severity of the same basic abnormality in the development of the vertebral centrum (Chahoud and Paumgartten, unpublished results).

Provision of adequate detail in descriptive terminology, use of severity grading where appropriate and recognition of the fact that different terminology is often used to describe different manifestations of the same underlying mechanism, are important considerations in facilitating accurate classification. For a matter-of-fact approach, however, these recommendations need sometimes to be tempered by considerations as follows. Given that conditions of exposure to a chemical substance (e.g. toxicokinetics/timing of embryofetal development) and genetic background are not exactly the same for all fetuses within a litter or a group, it is not unexpected that there might be a range of different changes