Reproductive Toxicology 80 (2018) 44–48

Contents lists available at ScienceDirect

Reproductive Toxicology

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

Best practices for developmental toxicity assessment for classification and labeling

George Daston[a,*], Aldert Piersma[b], Leonello Attias[c], Manon Beekhuijzen[d], Connie Chen[e], Jennifer Foreman[f], Nina Hallmark[g], Isabelle Leconte[h]

[a] Procter & Gamble, Cincinnati, OH, USA
[b] RIVM, Bilthoven, Netherlands, and IRAS, University of Utrecht, The Netherlands
[c] Istituto Superiore di Sanità, Rome, Italy
[d] Charles River, 's-Hertogenbosch, The Netherlands
[e] HESI, Washington, DC, USA
[f] ExxonMobil Biomedical Sciences, Inc., Annandale, NJ, USA
[g] Bayer SAS, Sophia Antipolis, France
[h] Charles River Laboratories France Safety Assessment SAS, France

ARTICLE INFO

Keywords:
Classification and labeling
Developmental toxicity

ABSTRACT

Many chemicals are going through a hazard-based classification and labeling process in Europe. Because of the significant public health implications, the best science must be applied in assessing developmental toxicity data. The European Teratology Society and Health and Environmental Sciences Institute co-organized a workshop to consider best practices, including data quality and consistency, interpretation of developmental effects in the presence of maternal toxicity, human relevance of animal data, and limits of chemical classes. Recommendations included larger historical control databases, more pharmacokinetic studies in pregnant animals for dose setting and study interpretation, generation of mechanistic data to resolve questions about whether maternal toxicity is causative of developmental toxicity, and more rigorous specifications for what constitutes a chemical class. It is our hope that these recommendations will form the basis for subsequent consensus workshops and other scientific activities designed to improve the scientific robustness of data interpretation for classification and labeling.

1. Introduction

The European Union has a harmonized classification and labeling process (CLP; Regulation (EC) no. 1272/2008 on classification, labelling and packaging) to identify and regulate hazards, including hazards to reproduction and development. Because of recent legislation on chemical safety (Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) in particular), many chemicals are undergoing this hazard-based classification and labeling evaluation. Because of the importance of this program to all sectors (*e.g.* public health, regulatory authorities and the regulated chemical industries), the European Teratology Society (ETS) and the Health and Environmental Sciences Institute (HESI) co-organized a workshop in September 2017 to air views on best practices for developmental toxicity assessment, especially as science progresses. (Presentations can be found online: http://hesiglobal.org/event/clp-cmr/). The purpose of this paper is to summarize the workshop and to provide recommendations for further research, dialog and training that would improve the classification process.

There are two possible classification categories for reproductive hazards (Table 1); the choice of category is based whether the chemical clearly poses a hazard to humans. Category 1 is split into category 1A, for which there are human data indicating an adverse effect on development, and category, 1B, for chemicals that have shown an effect in animal models, and therefore are presumed to pose a human hazard. Category 2 is for chemicals that cause an adverse effect, but for which the level of evidence is insufficient to draw a definitive conclusion about their hazard potential. Chemicals not meeting the criteria for any of the categories are unclassified. It may be worth noting that effects on sexual function and fertility, and on development, are considered separately. In addition, effects on lactation are allocated to a separate hazard category. The ETS-HESI workshop was focused on categorization of effects on development and not reproduction or lactation.

Each chemical and its data set is different, making it necessary to

* Corresponding author.
E-mail address: daston.gp@pg.com (G. Daston).

https://doi.org/10.1016/j.reprotox.2018.05.001
Received 14 February 2018; Received in revised form 30 April 2018; Accepted 3 May 2018
Available online 16 May 2018
0890-6238/

Table 1
Classification categories and criteria (from ECHA [2]).

Category		Criteria
1	Known or presumed human reproductive toxicant	Substances are classified in Category 1 for reproductive toxicity when they are known to have produced an adverse effect on sexual function and fertility, or on development in humans or when there is evidence from animal studies, possibly supplemented with other information, to provide a strong presumption that the substance has the capacity to interfere with reproduction in humans. The classification of a substance is further distinguished on the basis of whether the evidence for classification is primarily from human data (Category 1A) or from animal data (Category 1B).
1A	Known human reproductive toxicant	The classification of a substance in this Category 1A is largely based on evidence from humans.
1B	Presumed human reproductive toxicant	The classification of a substance in this Category 1B is largely based on data from animal studies. Such data shall provide clear evidence of an adverse effect on sexual function and fertility or on development in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary nonspecific consequence of other toxic effects. However, when there is mechanistic information that raises doubt about the relevance of the effect for humans, classification in Category 2 may be more appropriate.
2	Suspected human reproductive toxicant	Substances are classified in Category 2 for reproductive toxicity when there is some evidence from humans or experimental animals, possibly supplemented with other information, of an adverse effect on sexual function and fertility, or on development, and where the evidence is not sufficiently convincing to place the substance in Category 1. If deficiencies in the study make the quality of evidence less convincing, Category 2 could be the more appropriate classification. Such effects shall have been observed in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary non-specific consequence of the other toxic effects.

consider the weight of the evidence, and to apply expert judgment. As a consequence, there can be a considerable gray area around the level of evidence distinguishing category 1B from category 2 chemicals, or category 2 chemicals and unclassified chemicals.

A great deal of effort has gone into crafting guidance to minimize these gray areas, but even with this guidance there is room for different interpretations. Furthermore, science does not stand still, and progress in understanding modes of action, *in vitro* methods, and even changes in protocols in traditional animal studies all pose challenges to interpretation. Because of this, it is important to use a weight-of-evidence approach, considering all the information available on the chemical. Elements of weight-of-evidence evaluation include consistency across studies, dose-responsiveness, presence of maternal toxicity, and how the effects compare to historical control data and more. Each of these was discussed at the workshop and will be briefly described here.

2. Consistency of data and data requirements

Very few chemicals have human data; therefore, the vast majority of classification decisions are based on data from animal toxicity studies, particularly the OECD 414 protocol. It is not unusual for a chemical to have been tested for prenatal developmental toxicity in more than one species, or by more than one route of exposure. These results, along with other data, such as a one- or two-generation study, repeated-dose toxicity studies, and even non- regulatory studies like pharmacokinetics, *in vitro* assays, and mode of action studies all need to be considered in a weight of evidence assessment. Ideally, the toxicity of the chemical should be consistent across studies; *i.e.*, the same types of effects, and similar potency and dose-response characteristics. If these are not consistent, it is important to understand the reason. The first aspect to consider is study quality. Poor quality studies (studies with high variability, inadequate statistical power, limited or no analytical characterization of the test material, indications of poor animal husbandry or animal handling (*e.g.*, dosing errors)) should be weighted less than higher quality studies, and in the case of inconsistent results should probably be set aside. Quantitative weight of evidence schemes for developmental toxicity have been proposed [1], but even qualitative assessment of the data set on a chemical is useful in determining the validity of individual studies, and in drawing conclusions on the likelihood of the chemical being a developmental hazard.

In addition to study quality, there may be other reasons for inconsistencies that are related to the intrinsic properties of the chemical. It is possible for one species or strain of animal to respond differently, or for results to be different between dosing routes, because of differences in pharmacokinetics or metabolism. In the absence of other information, a precautionary approach is taken, and it is assumed that the more sensitive species is relevant for human prediction. Pharmacokinetic data is important for interpreting developmental toxicity results; while it is usually generated for pharmaceutical and agricultural chemicals, it is usually absent for other chemicals. It is also much more likely to be available for rats than rabbits. When it is present, pharmacokinetic data can provide powerful information for understanding lack of linearity in dose-response, *e.g.*, because of saturation of elimination pathways. These data can indicate dose levels that should not be exceeded because the responses at levels higher than these will not be relevant for predicting human hazard. Existing CLP guidance [2] indicates that marked TK differences between humans and test animals may be a reason not to classify a chemical. These data are unlikely to be available for most chemicals, but *in vitro* data on metabolism in human-derived liver cells may be useful in addressing species-specificity.

3. Historical control databases

Structural abnormalities typically occur at low prevalence and often amid a spontaneous (and sometimes fluctuating) background. Understanding this spontaneous background is important in interpreting results. Confidence that a malformation is caused by treatment (not just random or non-specific) is heightened if the incidence and severity of the malformation increases with dose, or if the malformation is very rare in controls. Interpretation is difficult when a low incidence of malformations is observed in a single dose group. If it occurs at the low dose it might be dismissed as background noise (particularly if there are no other manifestations of developmental toxicity at higher dose levels), but if the occurrence is in the high dose group the possibility exists that the effect is treatment-related. Good historical control data can be extremely important in determining whether the malformation is relatively high in prevalence, and in determining the range of prevalance. The former provides information on the overall rate of occurrence, whereas the latter (range) provides information as to whether the rate of occurrence in a particular group could plausibly be attributable to background. Ideally, the historical control database should be as large as possible, because any given malformation is likely to be relatively uncommon. Therefore, it is important to have a large number of studies from which to draw conclusions. Much work has been done on harmonizing terminology for malformations and variations [3], so differences in terminology should not be much of a barrier in combining data sets going forward. Unfortunately, most labs simply offer their own historical control data, which represents only a fraction of the existing data. Since there are relatively few animal suppliers the origin of the animals is not a significant variable in background malformation rate, making it possible to combine data from different labs, although methods used for fetal evaluation may differ among labs and

be a source of variability. Therefore, if there are discrepancies between a lab's historical control database and the overall database, an explanation should be sought. It would be beneficial to the science for contract research labs and large industry labs to pool their historical control data. This has been done in the past [4]; another effort is being made by HESI to create an updated pooled data base.

4. Maternal toxicity

Interpreting adverse developmental data observed in the presence of maternal toxicity is one of the most difficult problems in determining whether a chemical is a developmental hazard. Testing guidelines mandate that the highest dose produce some maternal toxicity, in order to maximize the chance of detecting a developmental hazard if one exists. (Pharmacokinetic data can be used to improve dose selection but as noted above is not always available.) Developmental effects observed at maternally toxic dose levels may be due to a direct effect on the embryo, or to an effect on the dam that produces a secondary effect on the embryo; *i.e.*, it does not possess an intrinsic developmental hazard. Compounds that have no intrinsic developmental hazard and only produce adverse developmental effects as a consequence or significant maternal toxicity should not be the subject of classification. Unfortunately, distinguishing which interpretation is correct is generally not possible using only the data reported in a prenatal developmental toxicity study. The data on maternal health is limited to body weight gain and food consumption over the period of gestation, as well as cage-side observations at the time of dosing and shortly thereafter. It is rare that clinical or histopathology data are available for the mother. This is a minimal amount of information compared to what is collected in a repeated-dose toxicity study, which routinely evaluate organ weights and histopathology, hematology and clinical chemistry. Although it may not be cost-effective to include all of these in a developmental toxicity study design, it is possible to use the information from repeated-dose studies in a weight-of-evidence evaluation of maternal toxicity (keeping in mind, however, that pregnant animals may respond differently), and if these studies are done before the developmental studies to consider including relevant assessments of specific organ toxicity in the study design.

The CLP guidance on how much maternal toxicity is too much is not specific, except in the case of maternal mortality. The guidance states that 10% or more mortality is excessive. Some clinical/behavioral manifestations of toxicity are mentioned in the guidelines but are limited: coma, hyperactivity, ataxia, labored breathing are mentioned as indicating excessive toxicity, but the list is clearly not exhaustive. Maternal weight gain is also addressed but there is no bright line provided as to how much weight gain decrement is too much. Only one guideline, the testing guideline for developmental neurotoxicity (OECD 426) provides an upper tolerable limit of 10% for maternal weight gain decrement. Expert workshops were held on the subject of maternal toxicity in Europe and the US [5,6]. While not unanimous, the consensus opinion from these workshops is that at some point maternal weight gain decrement becomes too severe to support interpretation, and the consensus appeared to be that 20% during the dosing period was too much. However, given the lack of unanimity, this may be an area of interpretation that deserves greater attention from expert societies like ETS.

Evaluation of individual dam and litter data is valuable in determining whether there is a direct association between maternal and developmental toxicity. In a group of 20–25 rats or rabbits there will always be some that are more highly affected than others. If developmental toxicity is attributable to maternal toxicity, then the litters of these highly affected animals should be more profoundly affected, too. While this association does not prove that the developmental effects are secondary to the maternal effects, it does increase the likelihood, just as a lack of association would decrease the likelihood. That said, in the case of severe maternal toxicity in some animals in a dose group, it is possible that others in that group are also affected but in a way that is not measured with the determined parameters. Thus, individual animal response is another factor in a weight-of-evidence assessment. Definitive evidence that developmental effects are maternally mediated can be obtained by dedicated mechanistic studies There has been a fair amount of research carried out to identify maternally-mediated mechanisms of developmental toxicity, and a number have been identified. These include various forms of anemia, secondary zinc deficiency through the induction of metallothionein in the maternal liver, altered acid-base balance, and others [7,8]. This research shows that there are many means by which perturbations in maternal physiology leading to temporary loss of homeostasis can indirectly affect development. However, none of these mechanisms can be demonstrated using only the information typically collected in standard developmental toxicity studies. Specially designed studies are required. When these studies are carried out, they should be considered as part of the weight of evidence in determining whether a chemical is a developmental hazard.

In vitro models, particularly whole embryo culture, can be very useful in determining whether a chemical has a direct effect on development. If a chemical, added to the culture medium at concentrations equivalent to what is present in the maternal system *in vivo*, does not cause adverse effects *in vitro*, this makes a strong case that the chemical is not a direct embryotoxicant. Of course, this requires some knowledge of the kinetics and metabolism of the compound *in vivo* so that appropriate concentrations are used, and the metabolite(s) are either generated in the culture system or are tested separately. Examples of the use of whole embryo culture in mechanistic studies include studies to show that glycolic acid is the active metabolite of ethylene glycol [9], or demonstrating that several agents that induce maternal metallothionein have no direct effect on the embryo, but alter development by a reversible decrease in circulating zinc [10].

5. Use of human data

It seems reasonable that human data should take precedence over animal data in determining human reproductive hazard. However, there are instances in which the human data may not be strong enough to support a decision. The example of boric acid was shared at the workshop. Boric acid adversely affected sperm production and embryonic development in animal studies [11]. In contrast, epidemiology studies on workers occupationally exposed to boron showed no effect on fertility. However, human studies did not provide sufficient evidence for a lack of boron-mediated effects on male fertility because of issues of statistical power. Therefore, boric acid was classified as a reproductive toxicant. In cases where good quality human data indicate an effect on reproduction it should be used for classification. However, lack of a positive result in humans may or may not result in no classification, depending on the quality of the study and the overall weight of evidence. Furthermore, in this case the animal data provide a plausible and relevant mechanism for human effects.

EFSA [12] published an opinion on best practices for using human data to identify public health concerns. While their review was limited to pesticides, the conclusions are broadly applicable. These include more standardized use of meta-analysis and systematic review in a weight of evidence approach, as well as considering biological plausibility in establishing causality.

6. Class effects

Chemicals with similar structures tend to have similar toxicity profiles. This is because similar chemicals interact with the same molecular targets, leading to the same biological effects. Accordingly, weight of evidence schemes favor grouping chemicals together, particularly if one or more members of a group cause adverse reproductive effects. However, there are limits to how different chemical structures can be and still fit within the same grouping for the purpose of

identifying toxicity. Two examples were shared during the workshop, phthalates and azoles. Phthalate esters are relatively simple in structure: a phthalic acid molecule esterified with alcohols (usually alkyl alcohols) of varying chain length. Phthalate esters with backbone chain lengths between four and six carbons in length consistently produce effects on male reproductive system development in rats, by interfering with fetal testosterone synthesis and by altering the expression of *insl*-1. Some phthalate esters with a three carbon backbone chain length produce these effects, but it is not consistent (*i.e.* disobutyl has effects whereas di-n-propyl effects are limited) [13]. However, phthalates outside this range, like dimethyl, diethyl, diheptyl, or dioctyl phthalate, have no effect on the reproductive system [14,15], and evaluations of gene expression in the fetal rat testis show different patterns of gene expression for these phthalates when compared to those that do affect development [16]. This example clearly shows that chemical classes are finite, and that, although some chemicals with very similar structures (*e.g.*, dibutyl and dihexyl phthalate) have the same toxicity, equally similar structures do not have the same toxicity (*e.g.*, dibutyl and diethyl phthalate). Conversely it should not be assumed that "activity cliffs" are the norm; where no biological activity has been identified or where structure activity relationships have been established according to read across guidelines, infinite testing would be required to establish a cliff does not exist which is clearly impractical and unhelpful.

Azole fungicides were also presented as an example of the limitations of considering all members of a chemical class to be equal. Azoles are complex molecules that share some common structural features and biological activity against CYP 51, an enzyme in the cholesterol synthesis pathway of fungi. Most azoles are also an inhibitor of CYP 19 (aromatase) in mammals, and possibly other molecular targets in the retinoic acid pathway. Despite the apparent similarity in molecular targets, the outcome of developmental toxicity studies can be very different, depending on the azole. Some produce a high rate of malformations (especially cleft palate) and post-implantation losses, while others produce only low rates of these effects, within or near historical control levels. Many of the azoles that have been through the CLH process have been classified as R1b, but recently an azole that had only limited developmental toxicity, in the presence of maternal toxicity, was classified as R2. Because of the complexity of the chemical structure of azoles, the pharmacokinetic profile, as well as affinity for CYPs, is likely to vary significantly among chemicals in this large group. Therefore, it is important that each chemical is assessed according to its specific data set and only if a classification decision cannot be reached, should additional information from other chemicals with a similar structure or hazard profile be considered.

7. Recommendations

There are a number of recommendations that the field of teratology should consider in the spirit of improving the classification process. These include:

- More comprehensive compilations of historical control data: Specific malformations don't always occur at high rates in a study, and therefore, comparison with historical data is important in interpreting whether a low level of malformations is consistent with the historical background. Testing is being done in a number of laboratories using the same strains of animals, and it should be possible to compile these data in a searchable manner. This was done in the 1990s by a regional teratology society (MARTA) and should be done on a continuing basis. HESI has plans to do this.
- More complete characterization of maternal toxicity: The typical evaluation of maternal toxicity is limited to mortality, macroscopic examination, body weight gain, food consumption and cageside observations. These may not be sufficient to fully characterize maternal toxicity. Additional endpoints that more sensitively evaluate maternal health, such as clinical pathology, organ weights or histopathology, should be added to study designs when these are affected in repeated dose toxicity studies. Perhaps these could be triggered based on the results of toxicity studies in non-pregnant animals, or of screening-level studies (*e.g.*, OECD 422).
- More routine use of pharmacokinetics: Pharmacokinetic data can be very useful in setting dose levels, and in interpreting study results, particularly when elimination pathways become saturated. Pharmacokinetic studies in pregnant animals are becoming increasingly routine in new drug evaluation and for agrochemicals. Guidance should be developed for when pharmacokinetic evaluations should be added to developmental toxicity studies.
- Guidance on how to use specialized studies to identify maternally-mediated modes of action: Non-guideline studies can be very powerful in providing data on whether developmental effects observed in the presence of maternal toxicity are secondary to that toxicity. An increasing number of technologies, from whole embryo culture to high throughput assay data (such as ToxCast) to toxicogenomics are available and can be applied to this problem. A workshop providing case studies for which these data have been generated would be valuable in developing guidance, and in training on how to interpret these data.
- Guidance on the limits of chemical classes: Chemical class information is an important aspect of weight of evidence; however, chemical classes are finite, and more guidance on how to identify the boundaries of a class would be useful to assessors. Some expert rules have been developed in the context of read-across [17,18]. These should be further developed and evaluated for their utility.

8. Conclusions

Classification is an important part of the European regulatory scheme for protecting the public from chemical hazards including those that affect reproduction and development. Therefore, it is essential that the process be scientifically robust. The ETS-HESI workshop was intended to review some of the difficult aspects of data interpretation as it pertains to classification of potential developmental toxicants and to offer suggestions. A series of recommendations is offered as a way of improving the process of developmental toxicity study design and interpretation.

References

[1] W. Dekant, J. Bridges, A quantitative weight of evidence methodology for the assessment of reproductive and developmental toxicity and its application for classification and labeling of chemicals, Regul. Toxicol. Pharmacol. 82 (2016) 173–185.
[2] ECHA, Guidance on the Application of CLH Criteria, Version 5.0, (2017) https://echa.europa.eu/documents/10162/23036412/clp_en.pdf/58b5dc6d-ac2a-4910-9702-e9e1f5051cc5.
[3] R. Solecki, M. Rauch, A. Gall, J. Buschmann, R. Clark, A. Fuchs, H. Kan, V. Heinrich, R. Kellner, T.B. Knudsen, W. Li, S.L. Makris, Y. Ooshima, F. Paumgartten, A.H. Piersma, G. Schönfelder, M. Oelgeschläger, C. Schaefer, K. Shiota, B. Ulbrich, X. Ding, I. Chahoud, Continuing harmonization of terminology and innovations for methodologies in developmental toxicology: report of the 8th Berlin Workshop on Developmental Toxicity, 14–16 May 2014, Reprod. Toxicol. 57 (2015) 140–146.
[4] MARTA, Historical control data, in: R.D. Hood (Ed.), Handbook of Developmental Toxicology, CRC Press, Boca Raton, FL, 1996, pp. 713–733.
[5] B.K. Beyer, N. Chernoff, B.R. Danielsson, K. Davis-Bruno, W. Harrouk, R.D. Hood, G. Janer, U.W. Liminga, J.H. Kim, M. Rocca, J. Rogers, A.R. Scialli, ILSI/HESI maternal toxicity workshop summary: maternal toxicity and its impact on study design and data interpretation, Birth Defects Res. B Dev. Reprod. Toxicol. 92 (2011) 36–51.
[6] L.D. Wise, J. Buschmann, M.H. Feuston, J.E. Fisher, K.W. Hew, A.M. Hoberman, S.A. Lerman, Y. Ooshima, D.G. Stump, Embryo-fetal developmental toxicity study design for pharmaceuticals, Birth Defects Res. B Dev. Reprod. Toxicol. 86 (2009) 418–428.
[7] E. Carney, Maternal physiological disruption, in: R.J. Kavlock, G.P. Daston (Eds.), Drug Toxicity in Embryonic Development, 1, Springer, Berlin, 1997, pp. 573–594.
[8] G.P. Daston, Relationships between maternal and developmental toxicity, in: C.A. Kimmel, J. Buelke-Sam (Eds.), Developmental Toxicology, 2nd ed., Raven, 1994, pp. 189–212.
[9] E.W. Carney, B. Tornesi, D.A. Markham, R.J. Rasoulpour, N. Moore, Species-specificity of ethylene glycol-induced developmental toxicity: toxicokinetic and whole embryo culture studies in the rabbit, Birth Defects Res. B Dev. Reprod. Toxicol. 83

(2008) 573–581.
[10] M.W. Taubeneck, G.P. Daston, J.M. Rogers, C.L. Keen, Altered maternal zinc metabolism following exposure to diverse developmental toxicants, Reprod. Toxicol. 8 (1994) 25–40.
[11] J.A. Moore, An assessment of boric acid and borax using the IEHR evaluative process for assessing human developmental and reproductive toxicity of agents. Expert Scientific Committee, Reprod. Toxicol. 11 (1997) 123–160.
[12] EFSA Panel on Plant Protection Products and their Residues (PPR), Scientific Opinion of the PPR Panel on the follow-up of the findings of the External Scientific Report 'literature review of epidemiological studies linking exposure to pesticides and health effects', EFSA J. 15 (2017) 1–101.
[13] A.M. Saillenfait, A.C. Roudot, F. Gallissot, J.P. Sabaté, Prenatal developmental toxicity studies on di-n-heptyl and di-n-octyl phthalates in Sprague-Dawley rats, Reprod. Toxicol. 32 (2011) 268–276.
[14] L.E. Gray Jr[1], J. Ostby, J. Furr, M. Price, D.N. Veeramachaneni, L. Parks, Perinatal exposure to the phthalates DEHP, BBP, and DINP, but not DEP, DMP, or DOTP, alters sexual differentiation of the male rat, Toxicol. Sci. 58 (2000) 350–365.
[15] A.M. Saillenfait, A.C. Roudot, F. Gallissot, J.P. Sabaté, M.C. Chagnon, Developmental toxic potential of di-n-propyl phthalate administered orally to rats, J. Appl. Toxicol. 31 (2011) 36–44.
[16] K. Liu, K.P. Lehmann, M. Sar, S.S. Young, K.W. Gaido, Gene expression profiling following in utero exposure to phthalate esters reveals new gene targets in the etiology of testicular dysgenesis, Biol. Reprod. 73 (2005) 180–192.
[17] S. Wu, K. Blackburn, J. Amburgey, J. Jaworska, T. Federle, A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments, Regul. Toxicol. Pharmacol. 56 (2010) 67–81.
[18] S. Wu, J. Fisher, J. Naciff, M. Laufersweiler, C. Lester, G. Daston, K. Blackburn, Framework for identifying chemicals with structural features associated with the potential to act as developmental or reproductive toxicants, Chem. Res. Toxicol. 26 (2013) 1840–1861.