On the impact of second generation mating and offspring in multi-generation reproductive toxicity studies on classification and labelling of substances in Europe

Rorije, Emiel; Muller, André; Beekhuijzen, Manon E.W.; Hass, Ulla; Heinrich-Hirsch, Barbara; Paparella, Martin; Schenk, Erna; Ulbrich, Beate; Hakkert, Betty C.; Piersma, Aldert H.

Published in:
Regulatory Toxicology and Pharmacology

Link to article, DOI:
10.1016/j.yrtph.2011.08.005

Publication date:
2011

On the impact of second generation mating and offspring in multi-generation reproductive toxicity studies on Classification and Labelling of substances in Europe

E. Rorije[a,*], A. Muller[b], M. E.W. Beekhuijzen[c], U. Hass[d], B. Heinrich-Hirsch[e] M. Paparella[f], E. Schenk[g], B. Ulbrich[e], B.C. Hakkert[a], A.H. Piersma[g,h].

[a] Expertise Centre for Substances, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands;

[b] Centre for Substances and Integrated Risk Assessment, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands;

[c] NOTOX B.V., ‘s-Hertogenbosch, The Netherlands;

[d] Department of Toxicology and Risk Assessment, National Food Institute, Technical University of Denmark, 2860 Søborg, Denmark;

[e] Department 6: Safety of Substances and Preparations, Federal Institute for Risk Assessment (BfR), Berlin, Germany;

[f] Environment Agency Austria, Vienna, Austria;

[g] Laboratory for Health Protection Research, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands;

[h] Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, The Netherlands.

* corresponding author:
Emiel Rorije, Substance Expertise Center (SEC) Postvak 1, RIVM, P.O.Box 1, 3720 BA Bilthoven, The Netherlands. tel.nr. +31-30-2742134, fax.nr +31-30-2744401, e-mail: emiel.rorije@rivm.nl

ACCEPTED FOR PUBLICATION IN

“Regulatory Toxicology and Pharmacology”

dd. 18-8-2011

ABSTRACT

The possible impact on classification and labelling decisions of effects observed in second generation parental (P1) and offspring (F2) parameters in multi-generation studies was investigated. This was done for 50 substances classified as reproductive toxicants in Europe, for which a multi-generation study was available. The P1 and F2 effects were compared to parental (P0) and first generation offspring (F1) effects with regard to type of effect as well as incidence, magnitude and severity (IMS), at any dose level. For every study with unique P1/F2 effects, or differences in IMS, the influence of the P1/F2 findings on the classification decision was investigated. Unique P1/F2 generation findings did not play a crucial role in the classification decision of any of the 50 classified substances, except for fenarimol. This substance however provided abundant alerts on the basis of its endocrine activity and developmental neurotoxicity and would therefore also be expected to be identified as a developmental neurotoxicant in an Extended One-Generation Reproductive Toxicity Study (EOGRTS). These findings, in addition to the increased number of parameters analysed, increased statistical power and reduced animal use, provide strong further support for replacement of the classical two-generation reproductive toxicity study by the EOGRTS in regulatory reproductive toxicity assessment.

KEYWORDS: Multi-generation reproductive toxicity;
Extended One Generation Reproductive Toxicity Study (EOGRTS);
Test guideline;
Developmental toxicity;
Classification and Labelling (EU C&L).

1. INTRODUCTION

Reproductive toxicity hazard assessment of substances in Europe is based on all relevant toxicological information retrieved from studies ranging from repeat dose tests in adult animals to the two-generation reproductive toxicity study (OECD Test Guideline 416, (OECD, 2001)). The latter study design includes exposure of adult males and females before mating (P0), and continued exposure of dams throughout pregnancy and weaning, exposure of the first generation offspring (F1) throughout life, including their mating (P1) and reproduction into a second generation offspring (F2), which is terminated at weaning. This study is time-consuming, requires no less than 2600 animals, and is limited as to the number of parameters included and the number of animals assessed for each parameter.

The extended one generation reproductive toxicity study (EOGRTS) (Cooper et al., 2006; OECD, 2010) has an innovated study design that includes extensive additional end point determinations. Novel end points include reproductive and endocrine parameters as well as developmental immunotoxicity and developmental neurotoxicity parameters. In addition, end points are assessed in more offspring than in the classical multi-generation study (e.g. the OECD TG 416 two-generation reproductive toxicity study (OECD, 2001) whilst the mating of the second generation (P1) and the second generation offspring (F2) are omitted from the protocol, unless triggered in specific cases. This new EOGRTS protocol is expected to provide a higher level of scientific information and at the same time substantially reduces animal use when no second generation offspring is produced.

The EOGRTS has been suggested as a possible replacement of the OECD TG 416 study. Discussion has focused on the necessity of producing a second generation offspring. These studies are applied in risk assessment as well as in classification and labelling of substances, for both of which European legislation is in place (EU, 2006; EU, 2008b). For risk assessment , the impact of the second generation offspring was addressed in our previous publication (Piersma

et al., 2011). We produced a unique database containing 498 rat multi-generation studies with 438 substances, mainly pesticides and industrial chemicals. A retrospective analysis of risk assessment reports considering these studies showed that the impact of the second generation had been negligible. This implied that the production of a second generation offspring might be omitted without impacting risk assessment outcome, saving significant time, and reducing animal numbers from 2600 to 1400 animals per study. These advantages are even more significant in the light of the EU REACH legislation (EU, 2006), which requires extensive animal toxicity testing in the coming five years. The contribution of reproductive toxicity generation studies has been estimated to amount to around 35% of all animal testing in REACH (van der Jagt et al., 2004), and omission of the second generation as indicated would therefore reduce animal use in REACH by around 15%.

We have concluded that also for classification and labelling in Europe (ECHA, 2011; EU, 2008b) it is highly unlikely that the second generation offspring would contribute significantly (Piersma et al., 2011). This analysis was based on relative parameter sensitivity in terms of lowest observed adverse effect levels (LOAELs) as compared between generations. However, it has been argued that the nature and the incidence, magnitude, and severity of effects might play a significant role specifically in view of classification and labelling. Thus, in this view, although the same LOAEL might have been derived for the first and the second generation, if the nature of the effect, or its incidence, magnitude or severity would be judged as more serious in the second generation, it might lead to a higher classification level.

In this manuscript, we have addressed the impact of the second generation parental and offspring parameters on classification and labelling in Europe. We have used the multi-generation study database to select those substances which had both a multi-generation study and in addition had been classified and labelled for effects on fertility, development or lactation. We identified 50 substances in the database satisfying these criteria, relevant for this analysis. For these substances the public records of the EU Specialized Expert (SE) and Technical

Committee (TC) meetings were studied to assess the impact that the second generation had on classification and labelling. For those compounds for which such records could not be retrieved, we did our own assessment of the likelihood that the second generation in the study would have specifically impacted on the classification and labelling. This analysis shows that, except for a single case, effects observed in second generation mating and offspring did not impact the decision on classification and labelling for reproductive toxicity. Moreover, the single case where second generation mating and offspring effects appeared to be instrumental for classification would be identified without any doubt as a reproductive toxicant in an EOGRTS without second generation mating and offspring. Therefore, this analysis supports the replacement of the OECD Test Guideline 416 two-generation reproductive toxicity study (OECD, 2001) with the EOGRTS (OECD, 2010). This replacement is expected to allow at least the same level of scrutiny for both risk assessment and classification and labelling, and moreover, in view of increased parameter number and enhanced power of the EOGRTS, it is anticipated to increase the likelihood for reproductive toxicants to be detected. The significant reduction in time and animal use provides further advantages that are more than relevant in view of implementation of the REACH legislation in Europe.

2. METHODS

The multi-generation reproduction toxicity study database was developed as described in detail before (Piersma et al., 2011). Briefly, the USEPA ToxRefDB format (Martin et al., 2009) was used and its content was extended with the database generated by Janer et al. (2007) and additional studies. The final database contained 498 multi-generation studies covering 438 substances. The substance list of the database was matched with the EU Classification and Labelling compound list, Annex VI to Regulation (EC) No 1272/2008 (EU, 2008a). In the database, 50 substances were found to carry a classification for reproductive toxicity on Annex VI. For these substances the multi-generation studies in the database were analyzed in detail as

to the nature, magnitude and severity of adverse effects found in the different generations within multi-generation studies, irrespective of the dose level at which they occurred. Study reports were consulted where necessary and whenever possible, and the reports of the EU Specialized Experts (SE) and Technical Committee (TC) on Classification and Labelling were taken into account to address the possibility of a unique contribution of the second generation mating and offspring (P1/F2). Summaries of the SE and TC meetings were available until 2010 on the website of the former European Chemicals Bureau (http://ecb.jrc.ec.europa.eu/) and most of them are still available through the H-class database from the Nordic Council of Ministers (http://apps.kemi.se/hclass/).

Throughout this manuscript, reference is made to the EU C&L system as used during the period of time addressed. This entails classification as Cat.1 for proven human reproductive toxicants, as Cat.2 for substances that should be considered as reproductive toxicants for humans based on animal studies, and as Cat.3 where there is some evidence for reproductive toxicity from animal data, but where the evidence is insufficient for Cat.2. Cat.1 and 2 reproductive toxicants are labelled with risk phrase R60 for fertility effects, and R61 for developmental effects, and Cat.3 reproductive toxicants are labelled with R62 for fertility effects and R63 for developmental effects. In the GHS system (UN, 2007) which is currently being introduced in Europe, Cat.1, 2 and 3 are generally replaced by the new Cat.1a, 1b and 2, with some (minor) changes as to the criteria for these categories. As mentioned, this manuscript refers to the old EU C&L system as it is based on references in which the old EU classification scheme has been used throughout.

3. RESULTS

The multi-generation study database (Piersma et al., 2011) of 438 substances contained 50 substances that had an EU classification for fertility, development, and/or lactation. These substances are given in Table 1. The multi-generation study summaries for these substances were analysed in order to assess whether the P1/F2 generation showed different types of effects, or the same effects but at lower doses as compared to the P0/F1 generation. It appeared that for 24 substances at least one multi-generation study showed effects that had been scored uniquely in the P1/F2 generation (Table 1). For the remaining 26 substances, the effects found in the first and second generation mating and offspring was not different in nature or toxicological relevance as indicated by the study summaries. Therefore, we conclude that for these 26 classified substances, the second generation mating and offspring was not crucial for the classification given.

Of the 24 substances with specific effects noted in the P1/F2 generation, 5 had a Cat.2, R60 classification and 8 had a Cat.3, R62 classification for fertility. Most of these substances had an additional classification for development. Of the remaining 11 substances without a classification for fertility, 6 had a Cat.2, R61 classification and 4 had a Cat.3, R63 classification for development, the remaining substance had a lactation label (R64) only. In the following, the multi-generation studies and the retrospective assessment of these groups of substances are discussed with specific reference to the impact of the P1/F2 generation findings. The meeting summary records of EU Specialized Experts (SE) and/or the EU Technical Committee (TC) on Classification & Labelling were consulted wherever possible. Substances classified for fertility were assessed first, as the multi-generation study is the principal test to detect possible effects on fertility. Subsequently, the possible impact of the multi-generation studies on developmental classifications and lactational labels were also considered as the study outcomes may also impact on these classifications.

3.1 Five substances with a Cat.2, R60 classification for fertility

3.1.1 Benomyl

The unique P1/F2 effect seen in the multi-generation study (Mebus, 1991) for benomyl was delayed eye opening. In addition, there was an F2 bodyweight and a litter viability effect, although these only occurred together with general toxicity seen in the P0 and F1 generations. However, the classification of Cat.2, R60 for fertility proposed by the SE (SE-Benomyl, 2001) was based on effects on testes seen in repeated dose studies and effects on testes and fertility in mating studies. Therefore, this classification was not dependent on the F2 effects in the Mebus study. The classification for development with Cat.2, R61 proposed by the SE was based on the brain and eye malformations observed in several developmental studies using gavage exposure. This classification would not have changed by the F2 effects in the Mebus study.

3.1.2 Cadmium(II)chloride

TC summary records (TC-Cadmium, 1998a; 1998b; 1998c) show that the 3-generation study (Nagymajtenyi et al., 1997) was not available at the time when classification of cadmium chloride was considered. Therefore it could not have affected the classification decision. This study has been discussed in our previous analysis of multi-generation studies (Piersma et al., 2011) Several other studies showing effects on fertility and development provided the basis for classification as Cat.2, R60-61.

3.1.3 DEHP

The 2-generation study showing the new P1/F2 effect (Schilling et al., 2001) could not have been decisive for the classification of DEHP for fertility or development. This study was not available to the TC during their assessment in 1999 and 2000, when the substance was classified with Cat.2, R60 based on extensive atrophy in testes and ovaries observed in repeated dose studies, and effects on fertility as observed in mating studies. The 2-generation study was not informative at all for the classification decision. The classification with Cat.2, R61 was primarily based on the results of developmental studies.

3.1.4 Vinclozolin

Matsuura et al. (2005) reported a reduced fertility in the P1 but not in the P0 males. In two earlier two-generation studies, reduced fertility in the P0 and P1 generations was observed through reduced live birth indices (Hellwig, 1992; 1994; 2000). Moreover, Matsuura et al. (2005) found histopathological effects on adrenals, pituitary, testis, seminal vesicles, and ovaries in P0 and P1. In addition, T4 was dose-dependently decreased, anogenital distance was reduced, and preputial separation was delayed in F1 males. These additional findings would be sufficient for classification for fertility in Cat.2, R60. The TC report specifically mentions reduced fertility and reduced epidydimal weight in support of Cat.2, R60 (TC-Vinclozolin, 1998). The classification for development (Cat.2, R61) was derived from effects on anogenital distance and its derived apparent sex ratio and malformations of the male reproductive organs in the F1.

3.1.5 1-Bromopropane

Unique effects of 1-bromopropane in the 2-generation study (Stump, 2001a) were a non-significant reduction of the fertility in the P1/F2 mating at the low and mid dose and a reduction in number of implantation sites at the mid dose. In addition, at lower levels of exposure, dose-dependent reductions in the numbers of pups born were recorded in both the F0 and F1 generations and reductions in the relative weights of epididymides, seminal vesicles and prostate were also noted. According to the TC summary records (TC-Bromopropane, 2002) the non-significant reduction in fertility and reduced number of implantation sites at the low and mid dose after the P1 mating were not considered as unique effects. The classification with Cat.2, R60 was based on effects on reproductive organs in the repeated dose toxicity studies and on mating effects and other fertility parameters in the 2-generation study. These effects were already observed in the parental animals and in the P0 mating. Therefore, it can be concluded that the second mating in this 2-generation study played a minor role in the argumentation for the classification and its absence would not have affected the classification with Cat.2, R60. Furthermore, according to the TC summary records, the classification with

Cat.3, R63 is based on the developmental effects observed in the prenatal developmental toxicity study. Therefore, it can be concluded that the second mating in this 2-generation study has not affected the classification with Cat.2, R63.

3.2 Eight substances with a Cat.3, R62 classification for fertility

3.2.1 Molinate

In the second 2-generation study for molinate (Moxon, 1994) at the mid dose decreased litter sizes were observed in the F2. This effect was probably secondary to several sperm and ovary effects observed in the F1 which would likely have influenced litter size. At the highest dose an increased gestational interval was observed for the P1 generation, but in addition other parameters including live birth index, ovaries, sperm and testes were affected in the F1, indicating that the classification for fertility as Cat.3 R62 had not been dependent on the observations in the P1/F2. This classification was mainly based on testicular changes, specific sperm lesions and reduced fertility of parental rats in one-generation studies. Moreover, Cat.3, R62 was proposed considering that the mechanism of action as well as kinetic differences rendered the findings less relevant for humans (TC-Molinate, 2003).

3.2.2 Octamethylcyclotetrasiloxane

In the 2-generation study (Stump, 2001b) the P1 but not the P0 showed a decrease in mating performance and in the number of animals producing litters, both at the study Lowest Effect Level (LEL) and above. However, the F1 and F2 mean live litter size were also reduced at the study LEL. The F1 also displayed disturbed oestrous cycles at the study LEL and above, as well as an increased pituitary gland weight at the highest dose. The SE (September 2006, accepted by TC, September 2007) considered that exposure of female rats around the time of mating causes a dose related reduction of numbers of corpora lutea, implantation sites and litter sizes.

Inhibition of the LH surge and subsequent ovulation is the mode of action, which is relevant to the human. However, the mechanism leading to the inhibition of the LH surge is unknown. Cat.3, R62 was given considering major differences in the regulation of ovulation in the human as compared to the rat making the relevance to humans doubtful. Clearly, the observations in the F1 would have been sufficient for classification as Cat.3, R62 for fertility.

3.2.3 BBP

In the multi-generation study (Aso et al., 2005) with BBP, at the high dose a decrease in fertility index was observed in the P1. However, already at the low and mid dose a decrease in epididymis weight in the F1 was noted, as well as a decreased anogenital distance. An earlier multi-generation study (Tyl et al., 2004) showed missing testes in one F2 male pup at the study LEL whereas no testes effects were observed in other generations. At the high dose mating and fertility effects were noted in the P1 as well as a decreased litter size in the F2, both of which were not noted in the P0/F1 generation. Multiple organ and reproductive organ effects as well as developmental landmark effects were found in the F1. The Tyl 2-generation study was provided during the assessment of BBP by the TC. However, TC summary records (TC-BBP, 2002) show that classification with Cat.3, R62 was already decided before these studies were provided and therefore they did not affect this classification. The TC may have used but does not mention in its summary a somewhat older 2-generation study (Nagao et al., 2000). This study is mentioned in the EU Risk Assessment Report (EURAR) for BBP (ECB, 2007). In this study the NOAEL value for effects on the reproductive organs in males was based on atrophy of the testis, epididymis, and seminal vesicle in the F1 at 10 or 18 weeks of age, and reduced reproductive organ weights in the F1 generation were reported at the next higher dose. Therefore, also in this study the second generation offspring did not provide information that would have changed the classification based on information from the first generation offspring. The available summary records of the TC contained no information about the basis for the Cat.2, R61 classification for development. There is no mention of a 2-generation study as the basis for this classification.

3.2.4 DBP

The multi-generation reproduction toxicity study for DBP (Wine et al., 1997) showed uniquely effects on pup body weights in F2 at the low dose, whereas both F1 and F2 body weights were affected at higher doses. However, effects on F1 litter size were observed at all dose levels, showing a dose-effect relationship. At mid dose additionally degeneration of the seminiferous tubules was observed. At the high dose in addition to F2 pup body weight effects a P1 fertility effect was noted which was not seen in the P0. At high dose the F1 showed litter size effects together with F1 pup body weight, ovary, testes, epididymis, liver, prostate and sperm effects. Although the exact types of effect noted were somewhat different among generations, they are clearly related. The effects observed in the F1 were sufficient for classification in Cat.3, R62. This classification was actually given on the basis of reduced fertility in female mice in a 1-generation study in the presence of systemic toxicity (Lamb et al., 1987; Morrissey et al., 1989). The classification with Cat.2, R61 was based on a developmental study in rats with exposure during gestation and lactation in which male reproductive tract malformations in offspring were observed at doses without maternal toxicity (TC-DBP, 2000).

3.2.5 Nonylphenol

In the NTP 2-generation study (NTP, 1997), at mid and high dose the F2 showed a decreased ovary weight and decreased epididymal sperm density (10%), whereas in the F1 a reduced bodyweight gain, histopathological changes in the kidney and delayed vaginal opening were observed. The developmental retardation observed in all offspring generations were considered by the TC (TC-Nonylphenol, 2001) as the basis for Cat.3, R62-63 fertility and development classifications and therefore the P1/F2 generation was not crucial.

3.2.6 Piperazine

In the 2-generation study (Wood and Brooks, 1994), at mid dose (LEL) reduced litter sizes were observed in the F1 and F2, and delayed sexual maturation in the F1 as well as a reduced number of implantation sites in the P0. At the high dose the P1 uniquely showed a decrease in

the number of pregnancies. The latter finding is to be expected on the basis of the reduced number of implantation sites in the P0, and will therefore not have been crucial for the Cat.3, R62 fertility classification. Cat.3, R63 was given on the basis of an increase in cleft palate incidence in the presence of maternal toxicity in a rat prenatal developmental toxicity study (TC-Piperazine, 2004).

3.2.7 Tepraloxydim

In the 2-generation study for tepraloxydim (Hellwig et al., 1997) only decreases in food consumption, bodyweight and bodyweight gain are noted in all generations at the highest dose, except for a developmental landmark effect which is observed only in the F2; a delay in time to eye opening. The information provided by industry and the summary records of the meetings in which the classification was discussed were confidential and can therefore not be discussed here. Our conclusion from the confidential data is that without any doubt the P1/F2 generation findings in the 2-generation study did not provide crucial information for the classification of tepraloxydim for reproductive toxicity in Cat.3, R62-63.

3.2.8 Fenarimol

In the case of fenarimol two multi-generation studies are available from the same test laboratory (Hoffman et al., 1977; Markham et al., 1978), both showing a decreased fertility in the P1 but not the P0. Cross-over data showed that this reduced fertility was clearly male mediated. No histopathological effect was seen anywhere, including thee male reproductive organs. Altered male mating behaviour due to brain developmental effects, mediated by aromatase inhibition, might have caused this effect. In the SE assessment of these studies in June 1998, fenarimol was noted as an aromatase inhibitor. The SE considered aromatase inhibition less relevant for humans in view of lower sensitivity in man, and concluded on a Cat.3, R62 classification for fertility (SE-Fenarimol, 1998). The TC (TC-Fenarimol, 1999) accepted the SE proposal. Therefore, the reduced fertility in the P1 due to fenarimol dosing, which was not observed in the P0, represents a case where second generation mating and offspring has provided crucial

information for classification and labelling. The sexe-dependent effect on functional brain development as observed in the multi-generation studies was also the basis for a developmental toxicity classification with Cat.3, R63. Therefore, also in this developmental toxicity classification the P1/F2 generation played a crucial role. In the discussion section of this manuscript we will address this case in view of current hazard assessment methodologies and argue that under current regulation this substance is expected to be identified as a reproductive toxicant irrespective of P1/F2 findings. Data on lactational transfer of fenarimol inspired the R64 labelling. For this finding a single offspring generation would obviously suffice.

3.3 Six substances with a Cat.2, R61 classification for development, without classification for fertility

3.3.1 DIHP

In the 2-generation study of DIHP (McKee et al., 2006), the unique F2 effect observed was a decrease in pup body weight at mid-dose. At the same dose level effects on P0 fertility and F1 sperm count were observed, as well as liver hypertrophy. This study was available when the TC (TC-DIHP, 2004) classified this substance for development in Cat.2, R61. However, there was already a majority opinion for classification with R61 before this study became available. The effects of the second generation did therefore not affect the classification for development.

3.3.2 Fluazifop-butyl

In a 2-generation study with fluazifop butyl (Willoughby et al., 1981), at high dose an effect on the fertility index was noted in the P1 but not in the P0. However, in the P0 also the number of implantations was reduced, gestational interval was increased, and in the F1 (but not in the F2) a decrease in live birth index was noted. Therefore, the unique F2 effects would not lead to a higher classification than the effects observed in the F1 and P0. Moreover, these effects

apparently did not lead to a classification for fertility. The classification with Cat.2, R61 for development was based on developmental effects in rat and rabbit in the developmental studies in the presence of no or slight maternal toxicity (TC-Fluazifop, 1998).

3.3.3 Flusilazole

The flusilazole multi-generation study (Pastoor et al., 1986) showed hydronephrosis in the F2b but not in F2a pups. This effect could also not be reproduced in a second study using identical dose levels (Mullin, 1990). The hydronephrosis in the first study was not dose-related and was considered by JMPR to be within the historical control range(Piersma et al., 2011). Having these studies available, the TC did not classify flusilazole for fertility (TC-Flusilazole, 1999). The Cat.2, R61 developmental toxicity classification was given on the basis of effects in rat and rabbit prenatal developmental toxicity studies. Therefore, the second generation mating and offspring in the multi-generation studies did not influence this classification.

3.3.4 Nickel(II)chloride

In the 2-generation study (RTI, 1988) increased malformed foetuses per litter (short rib) were observed in the F2a at the low dose, and an increased pup mortality and decreased live litter size were noted in the F1b. No other P1 or F2 effects were reported in this study, also not at higher doses. At the highest dose the effects which were only noted in the F1b were more pronounced (decreased litter size, increased pup mortality, decreased pup bodyweight) and seen in both cohorts F1a and F1b. The TC combined the results from several reproductive toxicity studies with different soluble nickel salts for the assessment of the classification of Nickel (II) chloride. The EURAR summary (ECB, 2008) explicitly noted that the malformations observed at low dose were not considered due to nickel because similar effects were not observed at higher doses. Therefore, the unique finding in the F2 did not impact on the decision about the classification of nickel chloride for developmental toxicity. In the absence of prenatal developmental toxicity studies, the classification for development (R61 Cat.2) was based on stillbirth and post-implantation / perinatal lethality observed in all generations in the 3-

generation study, in a 1-generation range-finding study and in a 2-generation study, at dose levels without maternal toxicity (TC-Nickel, 2004).

3.3.5 Nitrofen

In a 3-generation study (Ambrose et al., 1971) a slight body weight effect at the lowest effect level occurred in the P2 in breeding the F3, but not in the P1/F2 or P0/F1 generations (Piersma et al., 2011). This apparent decrease was due to a lower initial body weight of rats used for breeding and was considered not relevant for classification. The reason for classification with Cat.2, R61 by the TC could not be retrieved because this classification was already decided before 1997. However, the classification is most likely based on the well known teratogenic effects of nitrofen (Manson, 1986).

3.3.6 PFOS

In the 2-generation study (Luebker et al., 2005), at the highest dose level a gestational interval increase and a decrease in the number of implantation sites in the P1 but not the P0 was found. However, no F2 effects were reported. In addition, at the highest dose a decrease of the F1 live birth index and of the F1 viability index was noted, in the presence of body weight and food consumption effects. At one dose below the highest dose several developmental effects (delayed eye opening, delayed pinna unfolding, delayed development of surface righting ability) were observed in the F1, which could be relevant for classification as Cat.2, R61. According to the TC records (October 2006) the classification for developmental toxicity was based on litter resorptions in a mouse study. Clearly the second generation mating and offspring did not play a crucial role in the classification given.

3.4 Four substances with a Cat.3, R63 classification for development, without classification for fertility

3.4.1 Amitrole

The TC (October 2000) based their Cat.3, R63 classification on prenatal developmental toxicity studies as advised by the SE (SE-Amitrole, 2001). In their records, no mention is made of 2-generation studies or considerations about classification for fertility. One multi-generation study for amitrole (Gaines et al., 1973) showed a reduced number of litters in the F2a cohort at dose level 2 of 4 dose levels tested. Thyroid hyperplasia and reduced food consumption were noted at this dose level in the P0. At the next higher dose level, a reduced number of pups born and reduced pup survival were noted in the F1. There is a newer study (Richard, 1995) which used dose level 2 in the Gaines study as the high dose. In this study reduced number of implantation sites both in the P0 and P1, significantly decreased mean litter size at postnatal day 1 in the F1, and decreased F1 pup body weights during lactation were found, but also significantly decreased mating indices, decreased fertility indices, increased length of gestation in the P1 but not P0. These unique P1 fertility effects occurred at the same dose level and were of the same severity as the types of effects seen in the F1, and therefore the first offspring generation mating and the second generation would not affect the classification for developmental toxicity.

3.4.2 Fenpropimorph

Fenpropimorph was not classified as toxic to fertility, although the multi-generation study (Merkle et al., 1982) was available when the compound was assessed by the TC (TC-Fenpropimorph, 2003). The summary of the 2-generation study in rats provided in the classification proposal states that it showed some signs of developmental toxicity including increased number of stillbirths, decreased body weight gain in pups, and slight delays in physiological development of pups. The effects seen in the pups were considered most

pronounced in the F1 pups. The classification of Cat.3, R63 was primarily based on developmental toxicity observed in rabbit and rat prenatal toxicity studies. Therefore, also for this substance the second generation offspring in the 2-generation study did not provide information that would have changed the classification based on information from the first generation offspring.

3.4.3 Fentin hydroxide

In the multi-generation study for fentin hydroxide (Young, 1986) a slight F2 pup bodyweight effect was noted at mid-dose which was not noted in the F1. However, in the F1 decreased litter size and increase of several organ weights were noted at the same dose level. At the study high dose a decreased viability was observed in the F2 whereas in the F1 increased pup mortality and decreased litter size were noted. The F2 effects were of a similar severity as the F1 effects, and most likely did not play a decisive role in the classification of fentin hydroxide as developmental toxic, Cat.3, R63. This classification was probably based on the small increase in malformations in a prenatal developmental toxicity study with dermal exposure in rabbits (TC-Fentin, 2000).

3.4.4 Myclobutanil

At the high dose of the multi-generation study for myclobutanil (Brown, 1985) a small decrease in litter size in the F2 but not the F1 was noted. In both the F1 and the F2, reduced body weight, a decrease in the number of females delivering litters and an increased number of stillborn pups were observed. The unique F2 effect was observed at a dose level above the reproductive LOAEL, which was based on a minimal increase in proportion of dead pups in two matings of the P0. Therefore, the F2 effects were clearly not critical for the classification of myclobutanil for development, Cat.3, R63. This classification was based on an increase in embryotoxicity in a developmental study in the rat (TC-Myclobutanil, 1997).

3.5 One substance with a lactation label R64, without a classification for fertility and development

3.5.1 Lindane

In the multi-generation study for lindane (King, 1991), at the highest dose (one dose above the LOAEL for offspring effects according to JMPR) a delay in the onset of teeth and hair growth was observed in the F2, but not in the F1. These effects were seen together with reduced viability index and decreased body weight in both the F2 and the F1, and severe kidney and liver effects only in the F1. The specific developmental effect seen only in the F2 did not lead to classification of lindane for either fertility or developmental toxicity. It is unlikely that the F2 effects in this study played a decisive role in the labelling for lactation, R64. The classification with R64 was probably based on the presence of lindane in human milk and the results of a developmental neurotoxicity study (TC, April 2001)

4. DISCUSSION AND CONCLUSION

Our first multi-generation database analysis (Piersma et al., 2011) was founded on identifying differences between P0/F1 and P1/F2 generations on the basis of LOAEL comparisons between generations. LOAELs were defined as the lowest dose levels at which a toxicologically relevant adverse reproductive effect is observed, as decided by expert judgment. This expert assessment of toxicological relevance normally includes careful consideration of the nature and the incidence, magnitude and severity of the effects (IMS). On that presumption we argued that the LOAEL based analysis was also relevant for classification and labelling, as all relevant adverse toxicological effects (which includes their observed nature and IMS) should ultimately determine classification. The classification criteria specifically stipulate the importance of considering the nature and IMS of adverse effects in arriving at a decision on classification. In the current analysis we therefore took a step further, considering nature, magnitude and severity per generation in study reports in detail irrespective of the dose levels at which they occurred as well as revisiting the EU SE and TC summary records to establish how classification and labelling had been influenced by the P1/F2 generation findings in actual practice.

The European list of substances classified and labelled for reproductive toxicity (fertility, development and/or lactation) currently contains around 140 substances (EU, 2008a). For 50 of these a multi-generation study was available in our database (Piersma et al., 2011), of which 24 showed differences in findings as to type, incidence, magnitude and/or severity in the first versus second generation, 13 of which are classified (5 Cat.2 and 8 Cat.3) for fertility, which represents the foremost set of end points for which the multi-generation study is indicative. Of the 5 substances with a Cat.2, R60 classification for fertility, two (Cadmium and DEHP) did not have the multi-generation study available when the classification was decided, the remaining three showed relevant fertility effects already in the parental P0 generation (vinclozolin, 1-bromopropane, benomyl) and in repeat dose studies (benomyl). Seven of the 8 substances with a Cat.3, R62 classification for fertility (molinate, octamethylcyclotetrasiloxane, BBP, DBP,

nonylphenol, piperazine, tepraloxydim) showed fertility effects already in the P0/F1 sufficient for labelling. Among these, for BBP it is uncertain whether the multi-generation studies were available at the time and if they have been considered by the TC. For those substances among the 13 with a fertility classification that also carry a developmental toxicity classification, prenatal developmental toxicity studies were available or would clearly have sufficed to detect the developmental effect. Fenarimol provides the single exception in this analysis where the second generation was essential for classification and labelling, in Cat.3 for both fertility and development (see below).

Of the 24 substances which showed differences in findings in the first versus second generation, 10 are classified for developmental toxicity but not for fertility. Apparently, although multi-generation studies were available for these substances at the time of assessment by the TC, these provided no reason for classification for fertility. The developmental classification of these substances was based upon prenatal developmental toxicity studies, with the exception of nickel, which caused foetal and perinatal lethality in all generations in several generation studies. Therefore, for these 10 substances the second generation was not necessary for the classification given. The lactation label given to two substances could have been assigned using effects in the first generation offspring (fenarimol) or from a different type of study and the presence of the substance in human milk (lindane).

In the current analysis, fenarimol appeared as the single substance with a second generation specific effect specifically affecting classification and labelling. Fenarimol is an aromatase inhibitor, which affects the neonatal hypothalamus influencing subsequent expression of male sexual behaviour (Hirsch et al., 1987). This mechanism explains the presence of an observed effect on sexual behaviour in the P1 and not in the P0. The question remains whether modern test protocols without a second generation offspring would detect fenarimol. Clearly, *in vitro* assays for aromatase activity have identified fenarimol as a potent aromatase inhibitor (Vinggaard et al., 2000), providing an important alert for possible reproductive toxicity.

Andersen (Andersen et al., 2002) also found increased estrogen receptor (ER) transactivation and decreased androgen receptor (AR) transactivation by fenarimol in *in vitro* assays, which further stresses the alert for this substance. Furthermore, behavioural studies after prenatal or lactational exposure have also identified fenarimol as a neurodevelopmental toxicant. In the investigation by de Castro et al. (2007), lactational exposure resulted in a delayed climbing response, delayed righting reflex, delayed grip reflex, and persistent reductions in locomotion. According to these authors, the latter finding indicates persistent effects on the hypothalamic-pituitary-gonadal axis in the rat. These additional mechanistic investigations followed after the initial findings in the generation studies by Hoffman et al. (1977) and Markham et al. (1978), but such investigations are not among the European regulatory test requirements. However, in view of guinea pig data as well as data on human metabolism of testosterone available at the time, the SE and TC considered the mode of action of central aromatase inhibition observed in the rat as less relevant for man, and concluded that the risk of an effect on human fertility was low, leading to a Cat.3 classification rather than Cat.2 (SE-Fenarimol, 1998). The relevance of the estrogen signalling pathway for the sexual differentiation during brain maturation is not yet clear and considered of different significance for rodents versus primates including man (Li et al., 2008). The EOGRTS, proposed as a replacement for the 2-generation reproduction toxicity study in rats, is expected to detect the behavioural effects in the first generation offspring in its assessment of the functional observation battery and motor activity. Thus, the EOGRTS would be expected to detect fenarimol as a reproductive toxicant warranting classification and labelling without a second generation offspring being included. The case of fenarimol lends further support to the added value of the F1/P1 cohort for developmental neurotoxicity in the EOGRTS, and warrants its inclusion by default.

In conclusion, the present study reviewed actual practice of the official EU SE and TC with regard to classification and labelling where possible. We added our own detailed assessment of available data wherever differences in nature, incidence, magnitude and severity of effects between generations in the multi-generation reproduction toxicity study were observed. The

analysis showed that the P1/F2 generation findings in the 2-generation study, even if a unique F2 effect was observed, did not play a crucial role in the classification decision of any of the 50 classified substances for which a 2-generation study was available in our database, except in the single case of fenarimol. This substance however provides abundant alerts on the basis of its endocrine activity and is also expected to be identified detected as a developmental neurotoxicant in the EOGRTS. Moreover, the mechanism of action of this substance was considered by SE and TC as less relevant for humans. These findings strongly suggest that if the EOGRTS would have been used instead of the 2-generation reproduction study, conclusions of at least the same weight would have been drawn with respect to classification and labelling of the substances reviewed. As indicated in the OECD TG 443 on the EOGRTS, the accompanying draft guidance document will address particular aspects, e.g. the premating exposure duration period and the assessment of the developmental neurotoxicity and developmental immunotoxicity cohorts. Unique effects observed in the P1/F2 were mainly related to a reduction in fertility of the P1 due to changes in the reproductive organs as a result of the *in utero* exposure of the F1. Effects on reproductive organs can be detected both in one- and multiple generation studies as well as in repeat dose studies in adult animals. However, the EOGRTS is even more likely to detect such effects because of the detailed examination of the sexual organs of the F1 weanlings and adults. Therefore, in the absence of any crucial loss of data, and given the enhanced power and significantly increased number of end points relevant to reproduction and development, there is sufficient scientific justification to replace the two-generation reproductive toxicity study (OECD, 2001) with the new draft OECD EOGRTS guideline (OECD, 2010). This replacement provides a unique opportunity to significantly reduce animal use and testing time under the REACH legislation whilst at the same time enhancing scientific scrutiny. Given that the REACH legislation considers animal testing as a last resort and mentions the obligation to actively incorporate alternative methods wherever justified, this provides a formal basis for this change. In view of the legislative deadlines, an EU decision on the issue needs to take effect soon in order that REACH can effectively benefit from this replacement.

ACKNOWLEDGEMENTS

This project was sponsored by the Netherlands Ministry of Infrastructure & Environment, RIVM project REACH Supportive Research, M/601353. This financial support is gratefully acknowledged.

Conflict of Interest statement

The authors declare that there are no conflicts of interest.

REFERENCES

Ambrose, A. M., Larson, P. S., Borzelleca, J. F., Blackwell Smith, R., Hennigar, G. R., 1971. Toxicologic studies on 2,4-dichlorophenyl-p-nitrophenyl ether (TOK). Toxicol Appl Pharmacol. 19**,** 263-275.

Andersen, H. R., Vinggaard, A. M., Rasmussen, T. H., Gjermandsen, I. M., Bonefeld-Jorgensen, E. C., 2002. Effects of currently used pesticides in assays for estrogenicity, androgenicity, and aromatase activity in vitro. Toxicol Appl Pharmacol. 179**,** 1-12.

Aso, S., Ehara, H., Miyata, K., Hosyuyama, S., Shiraishi, K., Umano, T., Minobe, Y., 2005. A two-generation reproductive toxicity study of butyl benzyl phthalate in rats. J Toxicol Sci. 30 Spec No.**,** 39-58.

Brown, U., 1985. Two-Generation Reproduction Study with RH-53,866 Technical in Rats. Protocol No. 83P-364, Report No. 84R-117. Unpublished study prepared by Research Pathology Services, Inc. pp. 1-78.

Cooper, R. L., Lamb, J. C., Barlow, S. M., Bentley, K., Brady, A. M., Doerrer, N. G., Eisenbrandt, D. L., Fenner-Crisp, P. A., Hines, R. N., Irvine, L. F., Kimmel, C. A., Koeter, H., Li, A. A., Makris, S. L., Sheets, L. P., Speijers, G., Whitby, K. E., 2006. A tiered approach to life stages testing for agricultural chemical safety assessment. Crit Rev Toxicol. 36**,** 69-98.

de Castro, V. L., de Mello, M. A., Diniz, C., Morita, L., Zucchi, T., Poli, P., 2007. Neurodevelopmental effects of perinatal fenarimol exposure on rats. Reprod Toxicol. 23**,** 98-105.

ECB, 2007. EU Risk-Assessment Report Vol.76, 2007 on: benzyl butyl phthalate, CAS#: 85-68-7, EINECS#: 201-622-7. http://ecb.jrc.ec.europa.eu/DOCUMENTS/Existing-Chemicals/RISK_ASSESSMENT/REPORT/benzylbutylphthalatereport318.pdf, accessed May 2011.

ECB, 2008. EU Risk Assessment Report on: nickel, CAS#: 7440-02-0, EINECS#: 231-111-4. Final approved version awaiting publication. http://ecb.jrc.ec.europa.eu/documents/Existing-Chemicals/RISK_ASSESSMENT/REPORT/nickelreport311.pdf, accessed May 2011.

ECHA, 2011. Guidance on the Application of the CLP Criteria. Guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures. European Chemicals Agency, pp. 1-491. http://guidance.echa.europa.eu/docs/guidance_document/clp_en.pdf, accessed May 2011.

EU, 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union. L 396/1**,** p. 1–849.

EU, 2008a. Annex VI to Regulation (EC) No 1272/2008 on the classification, labelling and packaging of substances and mixtures. Harmonised classification and labelling for certain hazardous substances. Table 3.2 The list of harmonised classification and labelling of hazardous substances from Annex I to Directive 67/548/EEC.

EU, 2008b. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures. Official Journal of the European Union. L 353/1**,** p. 1-1355.

Gaines, T. B., Kimbrough, R. D., Linder, R. E., 1973. The toxicity of amitrole in the rat. Toxicol Appl Pharmacol. 26**,** 118-29.

Hellwig, J., 1992. Reproduction Study with Reg. No. 83 258 (Vinclozolin) in Rats Continuous Dietary Administration over 2 Generations (2 Litters in the First and 2 Litters in the Second Generation). Lab Project Number: 92/11251: 71R0375/88053. Unpublished report from BASF AG, Ludwigshafen, Germany.

Hellwig, J., Deckardt, K., Gembardt, C., 1997. Two-Generation Reproduction Toxicity Study in Wistar Rats: Continuous Dietary Administration. Unpublished Study Report Reg. No. 191819; Lab Project Number: 70R0268/92095F. BASF AG, Ludwigshafen, Germany.

Hellwig, J., et_al., 1994. Second reproduction study with Reg. No. 83 258 in rats. Continuous administration over 2 generations. Project No. 7080375/88119. Unpublished report from BASF AG, Ludwigshafen, Germany. Submitted to WHO by BASF AG.

Hellwig, J., van Ravenzwaay, B., Mayer, M., Gembardt, C., 2000. Pre- and postnatal oral toxicity of vinclozolin in Wistar and Long-Evans rats. Regul Toxicol Pharmacol. 32**,** 42-50.

Hirsch, K. S., Weaver, D. E., Black, L. J., Falcone, J. F., MacLusky, N. J., 1987. Inhibition of central nervous system aromatase activity: a mechanism for fenarimol-induced infertility in the male rat. Toxicol Appl Pharmacol. 91**,** 235-45.

Hoffman, D. G., Adams, E. R., Markham, J. K., Owen, N. V., Gibson, W. R., Morton, D. M., 1977. A multi-generation reproduction study with EL-222 in the rat. Unpublished report No.

R-715, R-1345, R-956, R-966 dated November 1977 Lilly Research Laboratories, USA. Submitted to WHO by DowElanco Europe, Wantage, Oxon, United Kingdom.

Janer, G., Hakkert, B. C., Slob, W., Vermeire, T., Piersma, A. H., 2007. A retrospective analysis of the two-generation study: what is the added value of the second generation? Reprod Toxicol. 24**,** 97-102.

King, V., 1991. Lindane: Reproductive Performance Study in Rats Treated Continously Through Two Successive Generations: Final Report (Addendum to MRID 422210). Lab Project Number: 91/0948: 91-CIL004-0948: CIL-004-LIND. Unpublished study prepared by Life Science Ltd.

Lamb, J. C., Chapin, R. E., Teague, J., Lawton, A. D., J.R., R., 1987. Reproductive effects of four phthalic acid esters in the mouse. Toxicol Appl Pharmacol. 88**,** 255-269.

Li, A. A., Baum, M. J., McIntosh, L. J., Day, M., Liu, F., Gray, L. E., Jr., 2008. Building a scientific framework for studying hormonal effects on behavior and on the development of the sexually dimorphic nervous system. Neurotoxicology. 29**,** 504-19.

Luebker, D. J., Case, M. T., York, R. G., Moore, J. A., Hansen, K. J., Butenhoff, J. L., 2005. Two-generation reproduction and cross-foster studies of perfluorooctanesulfonate (PFOS) in rats. Toxicology. 215**,** 126-48.

Manson, J. M., 1986. Mechanism of nitrofen teratogenesis. Environ Health Perspect. 70**,** 137-47.

Markham, J. K., Hoffman, D. G., Owen, N. V., Morton, D. M., 1978. A second multi-generation reproduction study with EL-222 (Lilly compound 56722) in the rat. Unpublished report No. R-636, R-1076, R-217 dated July 1978 from Lilly Research Laboratories, USA. Submitted to WHO by DowElanco Europe, Wantage, Oxon, United Kingdom.

Martin, M. T., Mendez, E., Corum, D. G., Judson, R. S., Kavlock, R. J., Rotroff, D. M., Dix, D. J., 2009. Profiling the reproductive toxicity of chemicals from multigeneration studies in the toxicity reference database. Toxicol Sci. 110**,** 181-90.

Matsuura, I., Saitoh, T., Ashina, M., Wako, Y., Iwata, H., Toyota, N., Ishizuka, Y., Namiki, M., Hoshino, N., Tsuchitani, M., 2005. Evaluation of a two-generation reproduction toxicity study adding endpoints to detect endocrine disrupting activity using vinclozolin. J Toxicol Sci. 30 Spec No.**,** 163-188.

McKee, R. H., Pavkov, K. L., Trimmer, G. W., Keller, L. H., Stump, D. G., 2006. An assessment of the potential developmental and reproductive toxicity of di-isoheptyl phthalate in rodents. Reprod Toxicol. 21**,** 241-52.

Mebus, C., 1991. Reproductive and Fertility Effects with DPX-T1991- 529 (Benomyl) Multigeneration Reproduction Study in Rats: Lab Project Number: 765-90. Unpublished study prepared by Haskell Lab, (E.I. du Pont de Nemours and Co.), pp. 1-1471.

Merkle, J., Freisberg, K. O., Hildebrand, B., Ruff, M., 1982. Report on a reproduction study with 4-[3-[4-(1,1-dimethylethyl)phenyl]-2-methyl]propyl-2,6(cis)-dimethylmorpholine (Reg. No. 108 406) in rats after oral administration (feeding). Two-generation study. Unpublished report dated 22 February 1982 from Department of Industrial Hygiene and Toxicology, BASF Aktiengesellschaft, Industrial Hygiene and Toxicology, BASF Aktiengesellschaft, Ludwigshafen, Germany. Submitted to WHO by Ciba-Geigy AG, Basel, Switzerland.

Morrissey, R. E., Lamb, J. C., Morris, R. W., Chapin, R. E., Gulati, D. K., Heindel, J. J., 1989. Results and evaluations of 48 continuous breeding reproduction studies conducted in mice. Fundam Appl Toxicol. 13**,** 747-777.

Moxon, M., 1994. ICIA5504: Multigeneration Study in the Rat: Lab Project Numbers: CTL/P/4213: 5504/83: RR0604. Vol. Unpublished study prepared by Zeneca Central Toxicology Lab, pp. 1-2127.

Mullin, L., 1990. Reproductive and Fertility Effects with Flusilazole Multigeneration Reproduction Study in Rats. Lab Project Number: HLR/424/90. Unpublished study prepared by E.I. du Pont de Nemours and Co., pp. 1-1199.

Nagao, T., Ohta, R., Marumo, H., Shindo, T., Yoshimura, S., Ono, H., 2000. Effect of butyl benzyl phthalate in Sprague-Dawley rats after gavage administration: a two-generation reproductive study. Reprod Toxicol. 14**,** 513-32.

Nagymajtenyi, L., Schulz, H., Desi, I., 1997. Behavioural and functional neurotoxicological changes caused by cadmium in a three-generational study in rats. Hum Exp Toxicol. 16**,** 691-9.

NTP, 1997. Nonylphenol: multigenerational reproductive effects in Sprague-Dawley rats when exposed to nonylphenol in the diet. R.O.W. Sciences study no. 8989-30. NTP Report # RACB94021. http://ntp.niehs.nih.gov/?objectid=071D9342-AE22-E76C-6C82D206806B591F, accessed May 2011.

OECD, 2001. Test No. 416: Two-Generation Reproduction Toxicity. OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects. http://www.oecd-ilibrary.org/environment/test-no-416-two-generation-reproduction-toxicity_9789264070868-en, accessed May 2011.

OECD, 2010. Draft Test Guideline for the Extended One Generation Reproductive Toxicity Study. OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects. http://www.oecd.org/dataoecd/23/10/46466062.pdf, accessed May 2011.

Pastoor, T. P., Rickard, R. W., Sykes, G. P., Kaplan, A. M., Gibson, J. R., 1986. Long-term feeding (combined chronic toxicity/oncogenicity study) and two-generation, four litter reproduction study in rats with INH-6573. Unpublished report No. HLR 32-86 from Haskell

Laboratory for Toxicology and Industrial Medicine, Wilmington, Delaware, USA. Submitted to WHO by E.I. du Pont de Nemours & Co., Inc., Wilmington, Delaware, USA.

Piersma, A. H., Rorije, E., Beekhuijzen, M. E., Cooper, R., Dix, D. J., Heinrich-Hirsch, B., Martin, M. T., Mendez, E., Muller, A., Paparella, M., Ramsingh, D., Reaves, E., Ridgway, P., Schenk, E., Stachiw, L., Ulbrich, B., Hakkert, B. C., 2011. Combined retrospective analysis of 498 rat multi-generation reproductive toxicity studies: On the impact of parameters related to F1 mating and F2 offspring. Reprod Toxicol. 31**,** 392-401.

Richard, J., 1995. Two generation study by oral route (dietary admixture) in rats. Unpublished report No. 6432 (study No. T5041262) from Centre International de Toxicologie, Evreux, France. Submitted to WHO by Bayer AG, Leverkusen, Germany.

RTI, 1988. Two generation reproduction study of nickel chloride administered to CD rats in drinking water: Fertility and reproductive performance of the P0 generation and F1 generation. Research Triangle Institute, USEPA, Washington, DC.

Schilling, K., Gembardt, C., Hellwig, J., 2001. Di-2-ethylhexyl phthalate - Two-generation reproduction toxicity study in Wistar rats. Continous dietary administration. Laboratory project identification 70R0491/97139. Unpublished. Experimental Toxicology and Ecology, BASF Aktiengesellschaft, D-67056 Ludwigshafen, Germany, pp. 1-1183.

SE-Amitrole, 2001. ECBI/59/00 - Add.1 - Rev.2, dd. 11.04.2001. SUMMARY RECORD Commission Group of Specialised Experts in the fields of Carcinogenicity, Mutagenicity and Reprotoxicity. Meeting in Arona, 06 - 08 September 2000. pp. 13-15. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924916, accessed May 2011.

SE-Benomyl, 2001. ECBI/59/00 - Add.1 - Rev.2, dd. 11.04.2001. SUMMARY RECORD Commission Group of Specialised Experts in the fields of Carcinogenicity, Mutagenicity and Reprotoxicity. Meeting in Arona, 06 - 08 September 2000. pp. 3-4. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924916, accessed May 2011.

SE-Fenarimol, 1998. ECBI/35/98 Rev.1, dd. 01.09.98. SUMMARY RECORD Commission Group of Specialised Experts in the fields of Carcinogenicity, Mutagenicity and Reprotoxicity. Meeting at Arona, 24-25 June 1998. pp. 1-3. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=923803, accessed May 2011.

Stump, D., 2001a. An inhalation two-generation reproductive toxicity study of 1-bromopropane in rats. Unpublished study No. WIL-380001. WIL Research Laboratories, Study sponsored by Brominated Solvents Committee (BSOC), Ashland (OH).

Stump, D., 2001b. A two-generation inhalation reproductive toxicity and developmental neurotoxicity study of octamethyltetrasiloxane (D4) in rats. Study No. 8713. Report No.2001-10000-50855, December 2001. WIL Research Laboratories,: Study sponsored by Dow Corning, Ashland (OH).

TC-BBP, 2002. ECBI/63/02 Rev.5, dd. 28.04.2003 SUMMARY RECORD Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances. Ispra, 25-27 September 2002. pp. 19-20. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=925054, accessed May 2011.

TC-Bromopropane, 2002. ECBI/42/02 Rev.2, dd. 25.11.2002. SUMMARY RECORD. Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances, Ljubljana, 15-17 May 2002. pp. 32-34. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924064, accessed May 2011.

TC-Cadmium, 1998a. ECBI/18/98 Rev.1, dd. 16.07.1998. SUMMARY RECORD. Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances, ECB Ispra, 6th - 8th May, 1998. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=922903, accessed May 2011.

TC-Cadmium, 1998b. ECBI/27/98 Rev.2, dd. 17.12.1998. SUMMARY RECORD. Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances, ECB Ispra, 1st - 3rd July 1998. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=923491, accessed May 2011.

TC-Cadmium, 1998c. ECBI/51/97 Rev.3, dd. 26.1.1998. SUMMARY RECORD. Commission Working Group on the Classification and Labelling of Dangerous Substances, Meeting at ECB Ispra, 15 - 21 October 1997. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924597, , accessed May 2011.

TC-DBP, 2000. ECBI/51/00 Rev.2, dd. 23.11.00. SUMMARY RECORD Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances, Hotel Concorde, Arona, 9 May 2000, ECB Ispra, 10 - 12 May 2000. pp. 52-53. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924586, accessed May 2011.

TC-DIHP, 2004. ECBI/147/04 Rev.3, dd. 08.11.2004. SUMMARY RECORD Meeting of the Technical Committee C&L on the Classification and Labelling of Dangerous Substances Parts I (existing substances) and II (general issues), Riga, 12-14 May 2004. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=922689, accessed May 2011.

TC-Fenarimol, 1999. ECBI/12/99 Rev.1, dd. 31.03.1999. SUMMARY RECORD Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances Pesticides. ECB Ispra, 18-20 November 1998. pp. 15-16. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=922443, accessed May 2011.

TC-Fenpropimorph, 2003. ECBI/52/03 Rev.1, dd. 20.10.2003. SUMMARY RECORD Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances. Pesticides – Health Effects. Ispra, 13-14 January 2003. pp. 2-3. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924615, accessed May 2011.

TC-Fentin, 2000. ECBI/43/99 Rev.2, dd. 01.03.2000. SUMMARY RECORD Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances. Pesticides. ECB Ispra, 19-21 May 1999. pp. 12-13. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924224, accessed May 2011.

TC-Fluazifop, 1998. ECBI/52/97 Rev.1, dd. 04.03.1998. SUMMARY RECORD Commission Working Group on the Classification and Labelling of Dangerous Substances. Pesticides. Meeting at ECB Ispra, 12 - 14th November, 1997 and ECBI/56/02 Add. 12, Rev. 6. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924625 and http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924841, accessed May 2011.

TC-Flusilazole, 1999. ECBI/12/99 Rev.1, dd. 31.03.1999. SUMMARY RECORD Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances. Pesticides. ECB Ispra, 18-20 November 1998. pp. 16-17. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=922443, accessed May 2011.

TC-Molinate, 2003. ECBI/56/02 - Add.1 - Rev.1, revised 7 Oct 03. Classification and labelling for adverse effects on fertility (CMR and Pesticides Working Groups). http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924823 , accessed May 2011.

TC-Myclobutanil, 1997. ECBI/27/97 Rev.3, SUMMARY RECORD Commission Working Group on the Classification and Labelling of Dangerous Substances. Pesticides. Meeting at ECB Ispra, 28 - 30 May, 1997. pp. 25-26. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=923487, accessed May 2011.

TC-Nickel, 2004. ECBI/139/04 Rev.2, dd. 28.10.2005. DRAFT SUMMARY RECORD Meeting of the Technical Committee C&L on the Classification and Labelling of Dangerous Substances. Arona, 21-24 September 2004. pp. 47-55. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=922513, accessed May 2011.

TC-Nonylphenol, 2001. SUMMARY RECORD Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances, ECB Ispra, 5 – 7 September 2001. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=924903, accessed May 2011.

TC-Piperazine, 2004. ECBI/147/04 Rev.3, dd. 08.11.2004. SUMMARY RECORD. Meeting of the Technical Committee C&L on the Classification and Labelling of Dangerous Substances Parts I (existing substances) and II (general issues). Riga, 12-14 May 2004. pp. 11-13. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=922689, accessed May 2011.

TC-Vinclozolin, 1998. SUMMARY RECORD. Meeting of the Commission Working Group on the Classification and Labelling of Dangerous Substances Pesticides, ECB Ispra, 18-20 November 1998. http://apps.kemi.se/hclass/DocumentDownload.aspx?DocId=922443, accessed May 2011, pp. 24-25.

Tyl, R. W., Myers, C. B., Marr, M. C., Fail, P. A., Seely, J. C., Brine, D. R., Barter, R. A., Butala, J. H., 2004. Reproductive toxicity evaluation of dietary butyl benzyl phthalate (BBP) in rats. Reprod Toxicol. 18**,** 241-64.

UN, 2007. Globally Harmonized System of Classification and Labelling of Chemicals (GHS). New York and Geneva. http://www.unece.org/trans/danger/publi/ghs/ghs_rev02/02files_e.html, accessed May 2011.

van der Jagt, K., Munn, S., Tørsløv, J., de Bruijn, J., 2004. Alternative approaches can reduce the use of test animals under REACH. European Commission, Joint Research Centre. Report EUR 21405 EN. http://ecb.jrc.ec.europa.eu/documents/REACH/PUBLICATIONS/Reducing_the_use_of_test_animals_under_REACH_IHCP_report.pdf, accessed May 2011.

Vinggaard, A. M., Hnida, C., Breinholt, V., Larsen, J. C., 2000. Screening of selected pesticides for inhibition of CYP19 aromatase activity in vitro. Toxicol In Vitro. 14**,** 227-34.

Willoughby, C. R., Tesh, J. M., Pritchard, A. L., et_al., 1981. Fluazifop Butyl: Effects Upon Reproductive Performance Of Rats Treated Continuously Throughout Two Generations. Unpublished Study Received Dec 4, 1981, LSR Report No. 81/Ily282/130.

Wine, R. N., Li, L. H., Barnes, L. H., Gulati, D. K., Chapin, R. E., 1997. Reproductive toxicity of di-n-butylphthalate in a continuous breeding protocol in Sprague-Dawley rats. Environ Health Perspect. 105**,** 102-7.

Wood, E., Brooks, P. N., 1994. Piperazine hydrochloride: Dietary two generation reproduction study in the rat. Report to Akzo Nobel from Safepharm Laboratories Ltd., Derby. (unpublished).

Young, D. L., 1986. A Dietary Two-generation Reproduction Study in Rats with Triphenyltin Hydroxide (Code: HOE 0.29664 OF ZD97 0004 technical substance): Final Report. Project No. WIL-39022 (A35378). Unpublished study prepared by WIL Research Laboratories Inc., Ashland, USA. Submitted to WHO by Hoechst AG., Frankfurt-am-Main, Germany, pp. 1-2815.

Table 1. 50 substances classified for reproductive toxicity (R60-63) and/or lactation (R64) in the combined retrospective database of multi-generation studies.

Substance	CAS number	Toxic to fertility		Toxic to development		Unique 2-gen effect in P1/F2?	P1/F2 effect critical to C&L?	Study summary source & Study reference where unique 2-gen effect is reported.
		Cat.2 R60	Cat.3 R62	Cat.2 R61	Cat.3 R63			
2-methoxyethanol	109-86-4	x		x		no		
Benomyl	17804-35-2	x		x		yes	no	JMPR (Mebus, 1991)
Boric acid	10043-35-3	x		x		no		
Cadmium(II) chloride	10108-64-2	x		x		yes	no	EURAR (Nagymajtenyi et al., 1997)
Carbendazim	10605-21-7	x		x		no		
DEHP	117-81-7	x		x		yes	no	EURAR (Schilling, 2001)
Potassium dichromate	7778-50-9	x		x		no		
Vinclozolin	50471-44-8	x		x		yes	no	JMPR (Matsuura et al., 2005)
1-Bromopropane	106-94-5	x			x	yes	no	CERHR (Stump, 2001a)
Glufosinate ammonium	77182-82-2	x			x	no		
2,4-dinitrotoluene	121-14-2		x			no		
2-hydroxyethyl picramic acid	99610-72-7		x			no		
Acrylamide	79-06-1		x			no		
Bisphenol A	80-05-7		x			no		
Molinate	2212-67-1		x			yes	no	EPA DER (Moxon, 1994)
Nitrobenzene	98-95-3		x			no	no	
Octamethylcyclotetrasiloxane	556-67-2		x			yes	no	SCCP (Stump, 2001b)
Azafenidin	68049-83-2		x	x		no		
BBP	85-68-7		x	x		yes	no	EURAR (Aso et al., 2005; Tyl et al., 2004)
DBP	84-74-2		x	x		yes	no	EURAR (Wine et al., 1997)
Dinoseb	88-85-7		x	x		no		
Linuron	330-55-2		x	x		no		
Carbon disulphide	75-15-0		x		x	no		
Nonylphenol	25154-52-3		x		x	yes	no	NTP (NTP, 1997)
Piperazine	110-85-0		x		x	yes	no	EURAR (Wood, 1994)
Tepraloxydim	149979-41-9		x		x	yes	no	EPA DER (Hellwig et al., 1997)
Fenarimol (+R64)	60168-88-9		x		x	yes	yes	EPA DER (Hoffman et al., 1977), JMPR (Markham et al., 1978)
DIHP	71888-89-6			x		yes	no	(McKee et al., 2006)
Dinocap	39300-45-3			x		no		
Ethylene thiourea	96-45-7			x		no		
Fluazifop butyl	69806-50-4			x		yes	no	EPA DER (Willoughby et al., 1981)
Flumioxazin	103361-09-7			x		no		
Flusilazole	85509-19-9			x		yes	no	JMPR (Pastoor et al., 1986), EPA DER (Mullin, 1990)
Isoxaflutole	141112-29-0			x		no		
Nickel(II)chloride	7718-54-9			x		yes	no	EURAR (RTI, 1988)
Nitrofen	1836-75-5			x		yes	no	JMPR (Ambrose et al., 1971)
PFOS	2795-39-3			x		yes	no	EPA DER (Luebker et al., 2005)
Tridemorph	24602-86-6			x		no		
Amitrole	61-82-5				x	yes	no	IUCLID (Gaines et al., 1973), JMPR (Richard, 1995)
Chlorotoluron	15545-48-9				x	no		
Cyproconazole	94361-06-5				x	no		
Fenpropimorph	67564-91-4				x	yes	no	JMPR (Merkle et al., 1982)
Fentin hydroxide	76-87-9				x	yes	no	EPA DER (Young, 1986)
Mancozeb	8018-01-7				x	no		
Maneb	12427-38-2				x	no		
Metconazole	125116-23-6				x	no		
Myclobutanil	88671-89-0				x	yes	no	EPA DER (Brown, 1985)
Tebuconazole	107534-96-3				x	no		
Toluene	108-88-3				x	no		
Lindane (+R64)	58-89-9					yes	no	EPA DER (King, 1991)