Comparison of rat and rabbit embryo–fetal developmental toxicity data for 379 pharmaceuticals: on the nature and severity of developmental effects

Peter T. Theunissen\textsuperscript{a,b,c}, Sonja Beken\textsuperscript{d}, Bruce K. Beyer\textsuperscript{e}, William J. Breslin\textsuperscript{f}, Gregg D. Cappon\textsuperscript{g}, Connie L. Chen\textsuperscript{h}, Gary Chmielewski\textsuperscript{i}, Luc De Schaepdrijver\textsuperscript{j}, Brian Enright\textsuperscript{k}, Jennifer E. Foreman\textsuperscript{l}, Wafa Harrouk\textsuperscript{m}, Alan M. Hoberman\textsuperscript{n}, Julia Y. Hui\textsuperscript{p}, Thomas B. Knudsen\textsuperscript{q}, Susan B. Laffan\textsuperscript{r}, Susan L. Makris\textsuperscript{s}, Matt Martin\textsuperscript{q}, Mary Ellen Mcnerney\textsuperscript{t}, Christine L. Siezen\textsuperscript{b}, Dinesh J. Stanislaus\textsuperscript{t}, Jane Stewart\textsuperscript{u}, Kary E. Thompson\textsuperscript{t}, Belen Tornesi\textsuperscript{k}, Jan Willem Van der Laan\textsuperscript{a,b}, Gerhard F. Weinbauer\textsuperscript{v}, Sandra Wood\textsuperscript{w}, Aldert H. Piersma\textsuperscript{a,x}

\textsuperscript{a}Centre for Health Protection, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
\textsuperscript{b}Medicines Evaluation Board, Utrecht, The Netherlands
\textsuperscript{c}Innovative Testing in Life Sciences and Chemistry, University of Applied Sciences Utrecht (HU), Utrecht, The Netherlands
\textsuperscript{d}Federal Agency for Medicines and Health Products, Brussels, Belgium
\textsuperscript{e}Sanofi U.S. Inc. Bridgewater, NJ, USA
\textsuperscript{f}Lilly Research Laboratories, Lilly Corporate Center, Indianapolis, IN, USA
\textsuperscript{g}Pfizer Worldwide Research & Development, Groton, CT, USA
\textsuperscript{h}ILSI-Health and Environmental Sciences Institute, Washington, DC, USA
\textsuperscript{i}Covance Laboratories Inc, Greenfield, IN, USA
\textsuperscript{j}Janssen R&D, Preclinical Development & Safety Beerse, Belgium
\textsuperscript{k}AbbVie Inc, North Chicago, IL, USA
\textsuperscript{l}ExxonMobil Biomedical Sciences, Inc, Annandale, NJ, USA
\textsuperscript{m}U.S. Food & Drug Administration, Silver Spring, MD, USA
\textsuperscript{n}Takeda Pharmaceutical Company, Deerfield, IL, USA
\textsuperscript{o}Charles-River Laboratories, Preclinical Services, Horsham, PA, USA
\textsuperscript{p}Celgene Corporation, Summit, NJ, USA

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

CONTACT Peter Theunissen P.Theunissen@cbg-meb.nl Medicines Evaluation Board, Graadt van Roggenweg 500, 3531 AH Utrecht, The Netherlands.
Abstract

Regulatory non-clinical safety testing of human pharmaceuticals typically requires embryo–fetal developmental toxicity (EFDT) testing in two species (one rodent and one non-rodent). The question has been raised whether under some conditions EFDT testing could be limited to one species, or whether the testing in a second species could be decided on a case-by-case basis. As part of a consortium initiative, we built and queried a database of 379 compounds with EFDT studies (in both rat and rabbit animal models) conducted for marketed and non-marketed pharmaceuticals for their potential for adverse developmental and maternal outcomes, including EFDT incidence and the nature and severity of adverse findings. Manifestation of EFDT in either one or both species was demonstrated for 282 compounds (74%). EFDT was detected in only one species (rat or rabbit) in almost a third (31%, 118 compounds), with 58% (68 compounds) of rat studies and 42% (50 compounds) of rabbit studies identifying an EFDT signal. For 24 compounds (6%), fetal malformations were observed in one species (rat or rabbit) in the absence of any EFDT in the second species. In general, growth retardation, fetal variations, and malformations were more prominent in the rat, whereas embryo–fetal death was observed more often in the rabbit. Discordance across species may be attributed to factors such as maternal toxicity, study design differences, pharmacokinetic differences, and pharmacologic relevance of species. The current analysis suggests that in general both species are equally sensitive on the basis of an overall EFDT LOAEL comparison, but selective EFDT toxicity in one species is not uncommon. Also, there appear to be species differences in the prevalence of various EFDT manifestations (i.e. embryo–fetal death, growth retardation, and dysmorphogenesis) between rat and rabbit, suggesting that the use of both species has a higher probability of detecting developmental toxicants than either one alone.

Keywords

Pharmaceutical testing; embryo–fetal developmental toxicity; cross-species evaluation; non-clinical
Introduction

Current regulatory guidelines for embryo–fetal developmental toxicity (EFDT) testing of human pharmaceuticals require studies in both a rodent and a non-rodent species, usually rat and rabbit (ICH S5(R2) 2005). This paradigm was developed in the mid-twentieth century in response to the pandemic of phocomelia associated with maternal exposure to thalidomide during early pregnancy since 1957 (Lenz 1961, 1988; McBride 1961; Schardein 2000). However, oral exposure to thalidomide did not result in phocomelia in the rat, whereas a syndrome of limb malformations was evident in the rabbit (Parkhie & Webb 1983; Shepard 1998; Schardein 2000). These observations suggested that the rabbit may serve as a better predictor than the rat to assess potential of EFDT in humans. However, dose ranges of thalidomide that were teratogenic in the rabbit induced embryo–fetal loss in the rat (Janer et al. 2008). This observation is consistent with current knowledge that the specificity of the manifestations of EFDT may vary greatly between species, and even between strains within the same species (Hurtt et al. 2003; Janer et al. 2008; Knudsen et al. 2009; Rorije et al. 2012).

This information raises questions about the need for, and timing of EFDT testing in a second animal species. (1) Does testing in a second species provide meaningful additional information for human risk assessment, or are there cases where a single species might be sufficient (i.e. clear teratogenicity with a biologically plausible mechanism in a relevant species)? (2) When developmental toxicity data are available in a pharmacologically relevant species at adequate exposures and clinical trials can maintain effective contraception, what impact would delaying the second species developmental toxicity study have on risk assessment and informed consent for the Phase III trials? The question of the need for, and timing of the second species EFDT study was discussed at several workshops (Brown et al. 2012; van der Laan et al. 2012). While investigations comparing rat and rabbit EFDT studies have been carried out in the fields of chemicals, pesticides, and veterinary drugs (Hurtt et al. 2003; Janer et al. 2008; Knudsen et al. 2009; Rorije et al. 2012), the aforementioned workshops discussed the need for a retrospective comparison of EFDT studies specifically conducted with pharmaceutical compounds.

To address the added value of the second species to the assessment of embryo–fetal toxicity of pharmaceuticals, a comprehensive database of rat and rabbit developmental toxicity studies was built through a collaboration between the Dutch National Institute for Public Health and the Environment (RIVM), the International Life Sciences Institute-(ILSI) Health and Environmental Sciences Institute (HESI) Developmental and Reproductive Toxicology (DART) Technical Committee, the Dutch Medicines Evaluation Board (CBG-MEB) and the US Environmental Protection Agency (US EPA). The database structure of the US EPA ToxRefDB (Knudsen et al. 2009) was used and extended to additionally collect pharmacokinetic information, such as maximum observed systemic concentration ($C_{\text{max}}$) and area under the concentration-time curve (AUC). This database provides a novel resource for retrospective comparison of rat and rabbit developmental toxicity studies. Unlike the previous comparisons of rat and rabbit developmental toxicity studies involving chemicals, pesticides and veterinary drugs (Hurtt et al. 2003; Janer et al. 2008; Knudsen et al. 2009; Rorije et al. 2012), this database was restricted to collected information on human
pharmaceuticals, including toxicokinetic data to allow consideration of maternal systemic exposure.

Using this database, we performed a retrospective analysis of the relative sensitivity of rat and rabbit along two main lines of investigation. In the present manuscript, a cross-species comparison is presented on the basis of the nature and severity of EFDT identified relative to the administered maternal dose levels. A second companion manuscript presents a cross-species analysis made on the basis of systemic exposure comparison and human equivalent dose conversion (Theunissen et al. 2016). A third manuscript will reflect on the implications for human health risk assessment (Cappon et al. in preparation).

**Methods**

**Data sources**

Information on rat and rabbit developmental toxicity studies was collected for human pharmaceuticals. Eligible studies met ICH S5(R2) guideline requirements, and quantitative data (e.g. Common Technical Document (CTD) tables or equivalently informative tabulated data from study reports) were used to populate the database. Written summaries were also used as a supplementary source of information. Data were acquired through two different sources: (1) EMA human medicinal product registration data were derived from the databases of the CBG-MEB in the Netherlands (including registered, withdrawn and rejected compounds that applied for marketing authorization between 2004 and 2012) and (2) Product data were acquired from a subset of HESI DART Technical Committee member companies. Products included registered human medicinal products, pharmaceutical compounds currently in development, and pharmaceutical compounds never submitted for regulatory approval. All HESI DART compounds in the database were anonymized by HESI staff, both as to chemical name/structure and company of origin and sent to RIVM for inclusion in the database. The HESI DART dataset was cross-referenced to the EMA dataset for duplicate compounds, which were later removed from the final database.

**The modified toxicology reference database**

All data were entered by RIVM staff using the format of US EPA’s “Toxicology Reference Database (ToxRefDB) stand-alone data entry tool” (Microsoft® Access based), enhanced to include toxicokinetic information. The EPA ToxRefDB database was originally designed for the ToxCast project to store data from regulatory toxicological animal test results (Martin et al. 2009). The ToxRefDB entry tool stores detailed study design, dosing, and observed treatment-related effects using a standardized vocabulary (Martin et al. 2009). The standardized vocabulary for developmental toxicity endpoints in the ToxRefDB is based on the DevTox lexicon (Makris et al. 2009), joined with standardized terms from the OECD-OPPTS vocabulary (OECD 2001) generating a thesaurus of 984 non-redundant terms applying to maternal and developmental endpoints in the original description of that database (Knudsen et al. 2009; Martin et al. 2009). The public version of ToxRefDB (http://www.epa.gov/nct/toxcast/data.html, release date December 2013) has information on prenatal developmental toxicity studies for approximately 600 chemicals that include 59 failed pharmaceutical compounds donated to the ToxCast program. These 59 compounds...
were not eligible for the present analysis due to lack of kinetic data for these entries. The enhanced database used for the present study will be made available, pending clearance of confidential data.

Effect type categorization

Developmental and maternal LOAELs were derived based on results in the individual EFDT studies. To obtain uniformity in the derivation of LOAEL among studies, maternal toxicity LOAEL (mLOAEL) and EFDT LOAEL (dLOAEL) were entered into the database after reexamination of the source data. The Dutch CBG-MEB and experts in the ILSI-HESI DART second species work group discussed ambiguous cases prior to inclusion in the data set. Ambiguities were scarce, and mostly included typographical errors and obvious misreads of the data. Critical effects determining the mLOAEL were related to adverse changes in maternal body weight (gain), food consumption, clinical signs, maternal gross pathology, maternal organ weight, or maternal mortality. Critical effects determining the dLOAEL related to adverse changes in fetal body weight, litter size, number of live and dead fetuses, resorptions, post-implantation loss, abortions, or fetal abnormalities (malformations and variations).

In order to compare the occurrence of the various manifestations of EFDT between species, findings defining EFDT were classified into four categories. These included: (1) embryo–fetal death (e.g. number of live/dead fetuses, decreased litter size, resorptions, post-implantation loss, and abortions), (2) growth retardation (e.g. fetal body weight reduction), (3) malformation (dysmorphogenesis), and (4) variation (changes over background incidence of morphological heterogeneity normally occurring within the species/strain). This fourfold grouping is a generally employed specification of classical manifestations of developmental toxicity. Structural anomalies in the ToxRefDB lexicon (n=984 terms) were categorized, where appropriate as malformations or variations based on the consensus of three HESI member companies’ current lexicons for fetal alterations (Bristol-Myers Squibb, Charles River Laboratories and GlaxoSmithKline). Malformation and embryo–fetal death were considered severe effects and were combined for some analyses, whereas growth retardation and variations were considered less severe manifestations of EFDT. Severity was defined operationally against the regulatory context, for which embryo–fetal death and malformations are usually interpreted as definitive manifestations of developmental toxicity, whereas growth retardation and variations, dependent on their incidence and magnitude may be considered less severe manifestations of developmental toxicity. This classification allowed the general comparative assessment as done in this study, but analysis of specific case comparisons would require more in depth assessment of available data.

Mode of action and therapeutic indication

Information on mode of action (MoA) and therapeutic indication, where available, was acquired through the European Public Assessment Reports for the EMA compounds (n=176 for MoA and indication) and through a survey for the HESI DART compounds (MoA n=28; indication n=42).
Database analysis and statistics

For analysis, relational data were extracted from the Microsoft Access based ToxRefDB standalone data entry tool using queries. This process is explained in more depth by Knudsen et al. (2009). For compounds with multiple studies in the same species, the study with the lowest LOAEL was used in the overall analysis. Information for each of the compounds and doses was extracted from the expanded ToxRefDB format and collected and processed in Microsoft Excel. Statistical analyses were performed in Graphpad Prism using a two-tailed Fisher’s exact test for contingency data, and findings were considered statistically significant at p<.05.

Results

Population of the database with developmental toxicity studies

Developmental toxicity studies of pharmaceutical compounds in both the rat and rabbit of acceptable quality (meeting ICH S5(R2) guideline requirements) submitted to EMA were entered into the enhanced ToxRefDB format. This resulted in studies for 176 compounds eligible for the database. Through the HESI DART second species working group pharmaceutical industry survey, nine companies provided data for 237 compounds. Eight of these compounds were duplicates of compounds submitted to EMA, and essential data were missing for 26 additional compounds, resulting in 203 compounds eligible for analysis. In total 379 compounds, comprising more than 800 studies were entered into the database. The mean and median of study year for the EMA compounds were similar for both the rat and rabbit (mean=1999.1; median=2000.0). Similarity of study timing between rat and rabbit studies was observed for industry survey compounds (mean rat=2005.7 and mean rabbit=2005.8; median rat and rabbit=2006.0). Studies acquired through the industry survey were on average more recently performed compared to the studies acquired through EMA, with a difference in median of approximately six years.

Frequency of different manifestations of developmental toxicity

Percentages of compounds causing any of the four categories of EFDT [malformation (M), embryo–fetal death (D), growth retardation (G) and variation (V)] were calculated for both the rat and rabbit for all 379 compounds present in the database (Figure 1). Occurrence of M, D, G and V at any dose in these studies is shown in Figure 1(A), whereas M, D, G and V observed at the dLOAEL are shown in Figure 1(B). There was no significant difference in frequencies of EFDT at any dose compared to EFTD at the dLOAEL only (cf. Figure 1(A) versus (B)). In total, 282/379 (74.4%) compounds induced some form of EFDT in at least one species. Comparing incidence of manifestations of EFDT between the rat and rabbit, embryo–fetal death (D) was significantly more often induced in the rabbit compared to rat (p<.001). Growth retardation and variations (G, V) were more frequently induced in the rat compared to the rabbit (p<.001). The percentage of compounds inducing malformations (M) was not significantly different between rat and rabbit (22.2% in rat and 20.6% in rabbit at any dose, 17.4% in rat and 12.9% in rabbit at dLOAEL only).

Co-occurrence of effects induced by each compound was evaluated. The number of compounds inducing malformations in both species (9.8% at any dose and 7.1% at dLOAEL...
Co-occurrence across the two species of any of the four effect categories or combinations thereof was limited. For instance, about 20% (rat 22.2%, rabbit 20.6%) of the compounds induced malformations at any dose studied in either rat or rabbit, (Figure 1(A)). Similar comparisons were observed for the two less-severe categories of EFDT, growth retardation and variations (Figure 1(A,B)).

**Frequency of malformation and/or embryo–fetal death**

The frequencies of malformations (M) and embryo–fetal death (D) were further investigated in detail. 143/379 (38%) and 170/379 (45%) of compounds induced M or D in the rat and rabbit, respectively (Figure 1(C)). At the dLOAEL only, M or D were induced by 120/379 (32%) of compounds in the rat and by 147/379 (39%) of compounds in the rabbit (Figure 1(D)). For this sub-selection of compounds, embryo–fetal death was significantly more often observed in the rabbit than in the rat (p<.001), whether considering data at any dose or dLOAEL only. Furthermore, compounds induced embryo–fetal death in the absence of malformations in the rabbit significantly more often than in the rat (p<.001). The percentage of compounds inducing malformations was higher in the rat than in the rabbit considering all doses, and still more so when considering the dLOAEL only (p<.001) compared to at any dose (p<.05). In addition, the percentage of compounds inducing malformations in the absence of embryo–fetal death was significantly higher in the rabbit than the rat (p<.05 at any dose; p<.001 at dLOAEL only). The percentage of compounds inducing both embryo–fetal death and malformations was comparable between rat and rabbit, both at any dose and at the dLOAEL, indicating that dLOAEL is representative of the cross-species differences across the full range of dosages tested.

Table 1 summarizes the comparison between rat and rabbit developmental toxicity manifestations based on the most severe category observed in each study. Overall, there was a 55.1% concordance between rat and rabbit. Differences in severity were roughly equally spread throughout the table, indicating similar overall sensitivities of developmental toxicity manifestations between rat and rabbit.

**Developmental toxicity in relation to maternal toxicity**

The values of dLOAEL as compared to mLOAEL were compared for all compounds, considering each of the different categories of EFDT noted at any dose (n=379; Figure 2). Selective EFDT (developmental effects at doses below maternal toxic doses; dLOAEL<mLOAEL) was observed for 59/379 (15.6%) compounds in the rat and 52/379 (13.7%) compounds in the rabbit. In total 91/379 (24%) compounds were selective developmental toxicants in at least one species, and 20/379 (5.3%) compounds were selective developmental toxicants in both rat and rabbit. Comparing incidences of each category of developmental toxicity between the rat and rabbit neither showed any significant difference for selective developmental toxicants (Figure 2, red), nor for compounds in which dLOAEL=mLOAEL (yellow) or dLOAEL>mLOAEL (green). Finally, the overall comparison between rat and rabbit of dLOAEL and mLOAEL for all 379 compounds showed no significant difference.
Comparing rat and rabbit strains

Frequency of the four different categories of EFDT was further investigated for each of the rat and rabbit strains used for the different studies in the database (Figure 3). Only two rat strains, Sprague–Dawley (SD) (n=332) and Wistar (n=43), were present in the database. For four compounds rat strain was not reported. Comparison of categories of developmental toxicity observed at any dose was not statistically significantly different from the distribution for all rat studies (Figure 3(A)). There were five rabbit strains present in the database (New Zealand White n=244; Dutch Belted n=43; Japanese White n=26; Himalayan n=25; Chinchilla n=4; one less defined group of Albino rabbits (Albino n=25); not reported rabbit strain n=12). Due to the low number of studies in Chinchilla and the unknown composition of the “not reported” groups for both the rat and rabbit, these compounds were not taken into account for further strain analysis. For rabbit strains, there was no significant difference in distribution compared to distribution of all rabbit studies for New Zealand White, Japanese White and Himalayan strains (Figure 3(B)). In the Dutch Belted strain, compounds induced significantly fewer malformations (p<.01), and significantly fewer cases with malformations and/or embryo–fetal death (p<.01). For the Albino rabbits, compounds induced significantly less embryo–fetal death. The large differences in the number of studies found across the various strains for either species do not allow any definitive conclusions about strain differences.

Organ specific malformations

Malformations were induced in the rat by 84/379 (22.2%) compounds and in rabbit by 78/379 (20.6%) compounds. Malformations observed in all studies at any dose were specified as to tissue type affected, and the number of studies in which they were observed was counted (Figure 4). No significant differences for specific malformations in any of the organs were observed between species. The greatest concordance between rats and rabbits seen in this analysis was malformations of bone (rat n=52, rabbit n=42), where in rat 17/52 (33%) and in rabbit 17/42 (40%) showed concordance of bone malformations (gray bars; Figure 4). This may be secondary to the fact that in EFDT studies, the skeletal system is assessed in much more detail than other tissues and organs. In addition, heart and great vessels received much attention in EFDT assessments, which is in line with the second highest concordance observed. Four of these 17 compounds were anti-cancer drugs inducing malformations in a wide variety of organs, in addition to other developmental effects such as embryo–fetal death. Two compounds were involved in iron metabolism and one compound was strontium based, which is known to have a specific effect on bone development (Shibata & Yamashita 2001). For the remaining 10 compounds, mode of action information was not available, thereby limiting further mode of action analysis. Cardiovascular anomalies (heart and/or vessels), which are among the most commonly occurring malformations in humans (CDC 2015), were noted at the same incidence in rats (20/379 compounds, 5.3%) and rabbits (19/379 compounds, 5.0%) with low concordance between species (6/379 compounds, 1.6%). Other interesting findings included the species-specific manifestation of abdominal malformations (“trunk” in Figure 4). In the rabbit (n=17), these consisted of omphalocele (n=6), gastroschisis (n=6), and umbilical hernia (n=5), whereas in the rat (n=12), omphalocele (n=1) and gastroschisis (n=0) were hardly observed, and a scatter of different trunk malformations was found. Only two compounds showed abdominal
malformations in both species, of which one also induced both scoliosis and a number of bone malformations in both species, and the other was an anti-cancer drug, inducing numerous different types of malformations.

**Developmental toxicity in one species only**

For 118/379 (31%) compounds, EFDT was observed in only one of the two species. In the dataset, 68/379 (18%) induced EFDT in the rat in the absence of any EFDT in the rabbit and 50/379 (13%) compounds induced EFDT in the rabbit in the absence of any EFDT in the rat (Figure 5). The percentage of rabbit-specific developmental toxicants inducing embryo–fetal death (41/50; 82%) was significantly higher than the percentage of rat-specific developmental toxicants inducing embryo–fetal death (26/68; 38%) (p<.001) (Figure 5(A)). For malformations, no such difference was observed (rat: 17/68, 25%; rabbit 7/50, 14%). For the species-specific developmental toxicants, there were statistically significantly fewer compounds inducing embryo–fetal death and malformation in rat only (33/68; 49%) compared to rabbit only (42/50; 84%) (p<.001). However, this difference might be attributed to the higher incidence of embryo–fetal death in rabbit as only 1/42 compounds induced malformations in the absence of embryo–fetal death in this species. For compounds causing EFDT in rat only, malformations were far more frequently observed in the absence of embryo–fetal death (9/33) than in the rabbit (1/42). The number of compounds inducing embryo–fetal death in the absence of malformations was lower in the rat (16/33) than in the rabbit (35/42). For the species-specific developmental toxicants, the occurrence of growth retardation and/or variation in the absence of malformations or embryo–fetal death was higher in the rat than in the rabbit (35/68 and 8/50 for rat and rabbit, respectively). It is possible that this disparity reflects lower numbers of evaluable rabbit fetuses, due to smaller litters and the higher incidence of embryo–fetal mortality in this species. Comparison of developmental toxicity versus maternal toxicity for those compounds inducing developmental toxicity in only one species neither showed significant difference between the rat and rabbit (Figure 5(B)), nor were these patterns significantly different compared to this distribution for all 379 compounds in the database (Figure 2).

Additional details available on organ-specific teratogens, compound mode of action, and a possible explanation for species difference of those compounds inducing malformations are shown in Table 2. No specific type of malformation or mode of action was obviously prevalent in either species for this group of compounds, but the dataset is too small to support any definitive inference.

**Discussion**

In the present analysis we evaluated the added value of two species EFDT studies carried out with pharmaceutical compounds, with a focus on differences in species sensitivity to EFDT, and the nature and severity of EFDT relative to the administered dose. We observed that in the rat, growth retardation and variations were more frequent, whereas in the rabbit embryo–fetal death was more often observed, with no significant difference in frequency of malformations between species. Among those studies that showed embryo–fetal death or malformations, rabbits more often showed embryo–fetal death, whereas rats more often showed malformations. The pattern of maternal toxicity (mLOAEL) relative to EFDT
(dLOAEL) did not differ significantly between these species. We found little evidence of strain-specific differences in overall EFDT categories in rats or rabbits. Organ-specificity of malformations was not different between species. For 31% of compounds, one species showed EFDT whereas the other did not, while 74% compounds showed EFDT in either one or both species. There was no observable relationship for species specificity, the presence of maternal toxicity, or pharmacological mode of action with a particular manifestation of EFDT.

Species sensitivity to EFDT

The frequency of EFDT in rats and rabbits was similar (61.2% versus 56.5%, Table 1). In total, 282/379 (74.5%) of compounds induced some manifestation of EFDT in at least 1 species. The percentage of compounds causing EFDT was ~60% and ~58% in rats and rabbits, respectively. Overall concordance for the presence or lack of EFDT between rat and rabbit was 69%. Concordance between rat and rabbit for the 282 compounds producing developmental toxicity was 58%; 164/282 compounds produced developmental toxicity on both species. Of the compounds for which rat and rabbit studies were not concordant, 18% (68/379) of the compounds produced EFDT in rats only and 13% (50/379) compounds produced EFDT in rabbits only. For the species-specific developmental toxicants, 25% (17/68) and 14% (7/50) in rats and rabbits, respectively, were selective developmental toxicants (i.e. EFDT occurred at doses not causing maternal toxicity). The manifestation of EFDT noted in the species-specific developmental toxicants was similar as observed for all 379 compounds in the database. Malformation was the manifestation of EFDT for the species-specific developmentally toxic compounds in 4% and 2% of the cases for rat and rabbit, respectively (17/379 of compounds in rat and 7/379 of compounds in rabbit produced malformation in one species but did not show any manifestation of EFDT in the other species). The occurrence of malformations in rabbit-specific developmental toxicants was observed in all cases except one accompanied by embryo–fetal death at the same dose levels.

In retrospect, there are numerous explanations for why an effect was only observed in one of the two species, such as differences in pharmacokinetics, differing profiles of maternal toxicity, or pharmacological response of the test species. The possible influence of kinetic parameters on species differences in this database is evaluated in much more detail in an accompanying report (Theunissen et al. 2016). The contribution of maternal toxicity to developmental outcome could explain some species sensitivity, especially in the rabbit where embryo–fetal wastage may occur secondary to maternal stress, e.g. through a period of gestational anorexia, which may perhaps in part explain these findings (Cappon et al. 2005). For some of these individual cases, we identified possible explanatory factors for these differences (Table 2). It is unknown whether the species specificity of the outcome of a developmental toxicity study may also relate to the pharmacological responsiveness of the test species. It is not known how frequently developmental toxicity of a pharmaceutical is linked directly to the intended primary pharmacology, but clearly there could be advantages to testing in species in which the pharmacological response, kinetics, and metabolism are known to be broadly similar to human.
**Manifestations of developmental toxicity**

As part of the US EPA ToxRefDB project, Knudsen et al. (2009) compared data from developmental toxicity studies performed in both rat and rabbit for 283 chemicals and pesticides. Although there are differences in analysis, design and compound classes (chemicals and pesticides versus pharmaceuticals), both analyses share similar findings. For instance, in both Knudsen et al. (2009) and our dataset, growth retardation and skeletal variations were more prevalent in the rat, whereas fetal losses (cf. embryo–fetal death in our study) occurred at a higher incidence in the rabbit. The similarity of these patterns among tested pharmaceuticals and among chemicals/pesticides suggests that the observed differences in frequency of the manifestation of developmental toxicity are likely not due to the nature of the compounds tested, but other factors which could be species-specific responses or differences in testing paradigms. For instance, rabbits tend to abort if they have periods of gestational anorexia, which may occur in the first days after the onset of dosing (Cappon et al. 2005). Knudsen et al. (2009) suggested an additional possible source of these differences. They hypothesized that the higher frequency of growth retardation and variations in the rat might be explained by differences in the dosing schedule (Knudsen et al. 2009). The post-treatment period from the end of exposure to terminal evaluation in their dataset was different between rat and rabbit (Knudsen et al. 2009). This was also the case in our dataset, where the average time lag between end of exposure and necropsy was 3.8 days for the rat, and 9.9 days for the rabbit. Therefore, reversible effects or effects associated with developmental delay (e.g. growth retardation and related variations, ossification effects) might be detected more easily in the rat than in the rabbit.

Organ specificity of malformations showed a similar pattern for both species (Figure 4). However, the number of compounds inducing a malformation in the same organ in both species was 40% or less. The highest overlap was observed for bone malformations (17 compounds out of 52 in rat and out of 42 in rabbit), the organ investigated most extensively, with the largest number of parameters in EFDT testing. An interesting case regarding organ-specific malformations was compound 10161 (Table 2), for which EFDT studies were performed in Sprague-Dawley rat, Wistar rat, and New Zealand White rabbit. Although for all three studies, dosing was performed within a similar dose range, no dLOAEL (indicating no developmental toxicity was observed at any dose in the study) was reached for the rabbit. This compound induced malformations in both rat strains with malformations of the eye (including microphthalmia and anophthalmia) in the Sprague–Dawley rat, and malformations of the paw (including brachydactyly, syndactyly and ectrodactyly) in the Wistar rat. Although comparisons such as this one need consideration of systemic dose as a possible confounder, this showed that even between strains of the same species, manifestation of malformations can be different.

**Role of maternal toxicity**

Janer et al. (2008), comparing rat and rabbit developmental toxicity studies of 117 chemicals, considered the presence of maternal toxicity as one of the possible factors that might influence differences in manifestation of developmental toxicity between species. In a subsequent study focusing on 22 known developmentally toxic chemicals in the rat, rabbit, or mouse, maternal toxicity was thought to play a role in inducing developmental
toxicity for 11/22 compounds (Rorije et al. 2012). In the current analysis, the overall numbers of selective developmental toxicants and toxicants with the mLOAEL at the same or higher dose as the dLOAEL was similar between rat and rabbit. This was observed for all compounds and for sub-selections of compounds, including: (1) compounds inducing malformation(s) and/or embryo–fetal death only; (2) compounds inducing effects in only one species; (3) between different species or strains; and (4) for each individual effect category. Findings of a similar nature were reported by Knudsen et al. (2009) for chemicals and pesticides, in which the effects on maternal body weight gain and concurrent fetal weight reductions appeared comparable between the rat and rabbit (Knudsen et al. 2009). In addition, the number of selective developmental toxicants was not significantly different between the Knudsen study (Knudsen et al. 2009) and our dataset [53/283 (18.7%) vs. 91/379 (24.0%), respectively; p=.10]. Based on the current analysis considering the entire dataset, no general conclusion could be drawn on the influence of maternal toxicity on differences in manifestation of developmental toxicity between the rat and rabbit. However, other investigations have suggested that the presence of maternal toxicity may be an explanatory factor in individual cases (Janer et al. 2008; Rorije et al. 2012).

**Strains within species**

Although the number of available Sprague–Dawley rat studies was much larger than the number of Wistar rat studies, both strains showed very similar frequencies of the four effect categories and distribution of fetal versus maternal toxicity (Figure 3(A)). The disparity in the number of compounds tested in the different rabbit strains (Figure 3(B)) makes any judgment about the performance of one particular rabbit strain versus another problematic. The different compounds tested in these strains may be a possible confounder.

**Developmental toxicity in one species only**

The most extreme species differences occur when compounds cause malformations in one species but do not show any manifestation of EFDT at all in the other species. In our database, this occurred for a small group of compounds (17/379 in rat, 7/379 in rabbit). For some of these individual cases, we identified possible explanatory factors for these differences (Table 2). Knowledge of compound pharmacology or MoA might explain differences for the antibiotic compound 10,104, which induced omphalocele in rabbit, but also abortion, embryo–fetal death, and maternal toxicity in the same species. Rabbits can be overly sensitive to antibiotics (ICH 2005), due to their gut flora being affected, which limits their use for testing EFDT of such compounds (Morris 1995). For two compounds, induction of severe maternal toxicity might be an explanation for the difference between species. For instance, compound 10033 induced maternal toxicity in the rabbit at a much lower systemic exposure than in the rat. The possible influence of kinetic parameters on species differences in this database is evaluated in much more detail in an accompanying report (Theunissen et al. 2016).

In addition, the higher frequency of growth retardation and variations in the rat may be partly explained by differences in time interval between last dosing and necropsy, as discussed earlier (Knudsen et al. 2009). The background incidence of manifestations of EFDT may also influence the significance level of dose-related EFDT. In addition, dose
spacing and choice of dose levels can influence differences in detection and manifestation of developmental toxicity between species. Besides these factors, repeating a developmental toxicity study in the same strain with the same compound and study design may result in dLOAELs differing by up to a factor of 10, as shown by Janer et al. (2008). This suggests that at least the same level of variability can be expected when comparing rat and rabbit studies, even without taking into account intrinsic species differences.

**Conclusion**

We performed a variety of analyses on the prevalence, nature and severity of developmental effects and their relation to administered dose and maternal toxicity after exposure to 379 pharmaceutical compounds in rat and rabbit regulatory EFDT studies. The question posed here was, does the second species provide meaningful additional information for human risk assessment, or are there cases where a single species might be sufficient, i.e. in the case of clear teratogenicity in a relevant species? This analysis, which is restricted to the analysis of the nature and severity of developmental toxicity at particular administered maternal doses and without taking into account pharmacodynamic response, kinetic exposure, or metabolism, underlines the utility of both species in assessing the developmental toxicity potential for any particular pharmaceutical. Finally, a companion report (Theunissen et al. 2016) considers this interspecies comparison from the perspective of systemic exposure, this aspect being crucial in view of comparison to intended human exposure and of therapeutic space. The overall picture emerging indicates that both species are similarly sensitive, although species differences were observed in the frequency of different EFDT manifestations. For 31% of compounds in the database, EFDT was observed in one species only, and 74% of compounds detected developmental toxicity in at least one species when two species were used. This suggests that the use of both species has a higher degree of detecting developmental toxicants than either one alone.

**Acknowledgements**

The authors gratefully acknowledge the extensive comments offered by the nine reviewers selected by the Editor and anonymous to the authors. These comments were extraordinarily helpful in revising the paper. The authors would like to thank the members of the HESI DART Technical Committee and the CBG-MEB and EMA for sponsoring the data and their scientific contribution, US EPA for providing and customizing the ToxRefDB, the University of Applied Sciences Utrecht for their administrative support and for hosting the project. HESI DART sponsor companies that contributed data toward this analysis were AstraZeneca, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Janssen R&D, Merck, Pfizer, Sanofi and Takeda.

Declaration of interest

HESI, a global branch of the International Life Sciences Institute, is a publicly supported, tax-exempt organization that provides an international forum to advance the understanding of scientific issues related to human health, toxicology, risk assessment and the environment through the engagement of scientists from academia, government, industry, NGOs, and other strategic partners. This HESI scientific initiative is primarily supported by in-kind contributions (from public and private sector participants) of time, expertise, and experimental effort. These contributions are supplemented by direct funding (that largely supports program infrastructure and management) that was provided by HESI’s corporate sponsors. A list of supporting organizations is available at www.hesiglobal.org.

The authors of the paper volunteered to serve on the HESI DART committee that conducted these analyses and prepared its conclusions and recommendations. The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of their employers, including companies that manufacture and market some of the chemicals evaluated or similar products, the Dutch CBG-Medicines Evaluation Board, the
Belgian Federal Agency for Medicines and Health Products, the U.S. Food & Drug Administration or the U.S. Environmental Protection Agency. There were no instances in which individuals advocated for particular treatment of a chemical manufactured and/or marketed by their employer. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. GDC is involved in birth defects litigation, but not as it pertains to species selection for developmental and reproductive toxicity testing. No other authors have appeared in any legal proceedings concerning the general topic of this paper within the last 5 years.

All studies were conducted in accordance with the relevant Policies on the Care, Welfare and Treatment of Laboratory Animals and were reviewed by the Institutional Animal Care and Use Committees where the work was conducted.

Funding

SLIM project, by the Dutch Government, Dept. Economical Affairs; The Utrecht Province and The Utrecht City Administration [PID101063]. ILSI HESI, 10.13039/100008663

Abbreviations:

| Abbreviation | Description |
|--------------|-------------|
| AUC          | area under the concentration-time curve |
| CBG-MEB      | Dutch Medicines Evaluation Board |
| C\text{max} | maximum concentration |
| CTD          | Common Technical Document |
| D            | embryo-fetal death |
| DART         | developmental and reproductive toxicology |
| dLOAEL       | lowest observable adverse effect level for developmental toxicity |
| EFDT         | embryo–fetal developmental toxicity |
| EMA          | European Medicines Agency |
| EPA          | Environmental Protection Agency |
| FDA          | Food and Drug Administration |
| G            | Growth Retardation |
| HESI         | Health and Environmental Sciences Institute |
| ICH          | International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use |
| ILSI         | International Life Sciences Institute |
| LOAEL        | lowest observable adverse effect level |
| M            | Malformation |
| mLOAEL       | lowest observable adverse effect level for maternal toxicity |
| MoA          | mode of action |
| OECD         | Organization of Economic Co-Operation and Development |
RIVM
National Institute for Public Health and the Environment, the Netherlands
ToxRefDB
Toxicology Reference Database
V
Variation

References

Brown ES, Jacobs A, Fitzpatrick S. 2012. Reproductive and developmental toxicity testing: from in vivo to in vitro. ALTEX. 29(3):333–339. [PubMed: 23019687]
Cappon GD, Fleeman TL, Chapin RE, Hurtt ME. 2005. Effects of feed restriction during organogenesis on embryo–fetal development in rabbit. Birth Defects Res Part B. 74:424–430.
Cappon GD, Theunissen PT, Beken S, Beyer BK, Breslin WJ, Chen CL, Chmielewski G, De Schaepdrijver L, Enright B, Foreman JE, et al. Comparison of rat and rabbit embryo-fetal developmental toxicity data for 379 pharmaceuticals: Implications for human risk assessment. (In preparation).
CDC. (2015). Congenital heart defects (CHDs): data & statistics. Atlanta, GA, USA. Available from: http://www.cdc.gov/ncbddd/heartdefects/data.html.
Hurtt ME, Cappon GD, Browning A. 2003. Proposal for a tiered approach to developmental toxicity testing for veterinary pharmaceutical products for food-producing animals. Food Chem Toxicol. 41:611–619. [PubMed: 12659713]
ICH. (2005). ICH harmonised tripartite guideline: detection of toxicity of reproduction for medicinal products & toxicity to male fertility S5(R2). Geneva, Switzerland. Available online: http://www.ich.org/products/guidelines/safety/safety-single/article/detection-of-toxicity-to-reproduction-for-medicinal-products-toxicity-to-male-fertility.html.
Janer G, Slob W, Hakkert BC, Vermeire T, Piersma AH. 2008. A retrospective analysis of developmental toxicity studies in rat and rabbit: what is the added value of the rabbit as an additional test species? Regul Toxicol Pharmacol. 50:206–217. [PubMed: 18171599]
Knudsen TB, Martin MT, Kavlock RJ, Judson RS, Dix DJ, Singh AV. 2009. Profiling the activity of environmental chemicals in prenatal developmental toxicity studies using the U.S. EPA’s ToxRefDB. Reprod Toxicol 28:209–219.
Lenz W. 1961. Kindliche missbildungen nach medikament-einnahme während der gravidität? Dtsch Med Wochenschr. 86:2555–2556.
Lenz W. 1988. A short history of thalidomide embryopathy. Teratology. 38:203–215. [PubMed: 3067415]
Makris SL, Solomon HM, Clark R, Shiota K, Barbellion S, Buschmann J, Ema M, Fujiwara M, Grote K, Hazelden KP, et al. 2009. Terminology of developmental abnormalities in common laboratory mammals (Version 2). Birth Defects Res Part B. 86:227–327.
Martin MT, Judson RS, Reif DM, Kavlock RJ, Dix DJ. 2009. Profiling chemicals based on chronic toxicity results from the U.S. EPA ToxRef Database. Environ Health Perspect. 117:392–399. [PubMed: 19337514]
McBride WG. 1961. Thalidomide and congenital anomalies. Lancet. 2:1358.
Morris TH. 1995. Antibiotic therapeutics in laboratory animals. Lab Anim. 29:16–36. [PubMed: 7707675]
OECD. (2001). OECD guideline for the testing of chemicals, no. 414: prenatal developmental toxicity study. Paris, France: Organisation for Economic Cooperation and Development.
Parkhie M, Webb M. 1983. Embryotoxicity and teratogenicity of thalidomide in rats. Teratology. 27(3):327–332. [PubMed: 6410529]
Rorije E, van Hienen FJ, Dang ZC, Hakkert BH, Vermeire T, Piersma AH. 2012. Relative parameter sensitivity in prenatal toxicity studies with substances classified as developmental toxicants. Reprod Toxicol. 34:284–290. [PubMed: 22652463]
Schardein J. 2000. Chemically induced birth defects. New York: Marcel Decker Inc.
Shepard TH. 1998. Catalog of teratogenic agents. Baltimore (MD): Johns Hopkins University Press.

Shibata S, Yamashita Y. 2001. An ultrastructural study of osteoclasts and chondroclasts in poorly calcified mandible induced by high doses of strontium diet to fetal mice. Ann Anat. 183:357–361. [PubMed: 11508362]

Theunissen PT, Beken S, Beyer BK, Breslin WJ, Gappon GD, Chen CL, Chmielewski G, De Schaepdrijver L, Enright B, Foreman JE, et al. 2016. Comparing rat and rabbit embryo-fetal developmental toxicity data for 379 pharmaceuticals: on systemic dose and developmental effects. Crit Rev Toxicol. [Epub ahead of print], doi: 10.1080/10408444.2016.1224808.

Van der Laan JW, Chapin RE, Haenen B, Jacobs AC, Piersma A. 2012. Testing strategies for embryo-fetal toxicity of human pharmaceuticals. Animal models vs. in vitro approaches: a workshop report. Regul Toxicol Pharmacol. 63(1):115–123. [PubMed: 22449444]
Figure 1.
Occurrence of embryo–fetal death (D), malformation (M), growth retardation (G) and variation (V) for all compounds (n=379) at (A) any dose and (B) observed at the dLOAEL (blue=rat, red=rabbit, gray=concordance of effect category induced by the same compound in both rat and rabbit). Occurrence of D and M for compounds inducing M and or D in rat and rabbit (C) at any dose and (D) specifically at the dLOAEL. (*p<.05; ***p<.001. Fisher’s exact test).
Figure 2.
dLOAEL in relation to mLOAEL for each effect category at any dose in rat (R) and rabbit (B) for all compounds \((n=379)\). No significant differences were observed between species for selective developmental toxicity \((d\text{LOAEL}<m\text{LOAEL}; \text{red})\) or developmental toxicity in presence of maternal toxicity \((\text{yellow} + \text{green})\). (Fisher’s exact \(t\)-test).
Figure 3.
Percentage of studies within (A) rat strains and (B) rabbit strains inducing a specific effect category. Embryo–fetal death (D), malformation (M), growth retardation (G) and variation (V). (*p<.05; **p<.01. Fisher’s exact test).
Figure 4.
Number of compounds inducing malformations specified for each organ for rat (22.2% of the compounds; blue) and rabbit (20.6% of the compounds; red) at any dose. Gray bars represent number of compounds inducing a malformation in the same organ in both rat and rabbit. No statistically significant differences were noted between species in occurrence of organ specific malformations (Fisher’s exact test).
Figure 5.
Percentage of compounds inducing developmental toxicity in one species and not in the other species. Rat dLOAEL, rabbit no dLOAEL \( n=68 \) (indicating no developmental toxicity was observed at the highest dose in the study); Rabbit dLOAEL, rat no dLOAEL \( n=50 \). (A) Occurrence of effects observed for these compounds in rat (blue bars) and rabbit (red bars). (B) Developmental toxicity in relation to maternal toxicity for this sub-set of compounds. Embryo–fetal death (D), malformation (M), growth retardation (G) and variation (V).
Table 1.

Comparison of compound induced effects between rat and rabbit.

| Rat     | NE   | G+V only | M+D | Total |
|---------|------|----------|-----|-------|
| NE      | 97   | 8        | 42  | 147   |
| G+V only| 35   | 19       | 35  | 89    |
| M+D     | 33   | 17       | 93  | 143   |
|         | 165  | 44       | 170 | 379   |

(A) Absolute numbers; (B) percentages; M + D: malformation and/or embryo–fetal death; G + V only: growth retardation and variation in the absence of malformation or embryo–fetal death; NE: no developmental effects observed.
| Compound Code | Organ specific malformations at any dose | MoA | Possible explanation for species difference |
|---------------|-----------------------------------------|-----|--------------------------------------------|
| 10030         | Hydronephrosis, umbilical hernia and dilated cerebral ventricle | antifibrotic | Rat dLOAEL after dietary exposure (dose not measured); 2nd rat gavage study no dLOAEL |
| 10033         | Bent rib | benzodiazep agonist | Maternal toxicity in rabbit at much lower doses than rat. |
| 10071         | Ventricular septal defect |  |  |
| 10072         | Malpositioned lumbar ribs | NK-2 Receptor Antagonist | TK: AUC and Cmax at dLOAEL much higher in rat than at highest dose tested in rabbit |
| 10076         | aortic arch, facial, lung and bone malformations | synaptic function | TK: AUC and Cmax at rat dLOAEL higher than at highest dose tested in rabbit |
| 10090         | Absent Sternebrae |  |  |
| 10095         | Microphthalmia, n = 1 hydrocephalus and situs inversus |  |  |
| 10108         | Absent Ribs | anti-Diabetic: DPP-4 inhibitor | Absent ribs in historical controls, dLOAEL due to rib variations without maternal tox |
| 10129         | Multiple bone, limb and facial malformations |  | Malformations and maternal tox in rat at lower AUC (4x) and Cmax (9x) compared to highest dose tested in rabbit |
| 10161         | Multiple paw/digit (Wistar) versus eye (SD) malformations | vasopressin2 antagonist | Two rat strains show different types of malformations |
| 10188         | Hydrocephalus, microphthalmia, anophthalmia | platinum-based antineoplastic | Reason for species difference not clear |
| 10211         | Absent sacral/caudal vertebrae |  |  |
| 10257         | Fused sternbrae |  |  |
| 10311         | Malformations, type or organ not specified | 5-HT4 agonist | High incidence of maternal toxicity (mortality) in rat at dLOAEL, no maternal toxicity in rabbit at a similar TK. 7.5x higher concentration in rat vs rabbit placenta |
| 10370         | Bent ribs | dehydroepiandrosteron (DHEA) | Reason for species difference not clear |
| 10389         | Ossification center in eye, misshapen lens |  | Reason for species difference not clear |
| 10426         | Fused bones | anti-Diabetic: SGLT2 inhibitor | Reason for species difference not clear |
| 10094         | Absent gall bladder, bent bones |  | TK: AUC in rabbit lower compared to highest dose tested in rat. Cmax is similar. Differences in ADME? |
| 10104         | Omphalocele | antibiotics: quinolone | Mode of action: rabbit is known to be oversensitive to antibiotics |
| 10137         | Retrocaudal ureter, levocardia | antifungal: echinocandin | Reason for species difference not clear |
| Compound Code | Organ specific malformations at any dose       | MoA                      | Possible explanation for species difference                                      |
|---------------|-----------------------------------------------|--------------------------|------------------------------------------------------------------------------------|
| 10224         | Small or absent cerebrum                      |                          | Reason for species difference not clear                                              |
| 10330         | Enlarged cerebral ventricle                   | HPG axis modulation      | Reason for species difference not clear                                              |
| 10436         | Fused maxilla, detached sternum               | inducer of cytotoxicity  | **Mode of action:** developmental toxicity is expected to be induced by a cytotoxic agent |

(A) Seventeen compounds induced malformations in rat with no dLOAEL in rabbit. (B) Seven compounds induce malformation in rabbit with no dLOAEL in rat. When available, mode of action and a possible explanation for the species differences in sensitivity for developmental toxicity are provided (given in bold).