Organisation for Economic Co-operation and Development

ENV/JM/MONO(2020)21

Unclassified English - Or. English

9 October 2020

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

Cancels & replaces the same document of 24 September 2020

Case study on the use of Integrated approaches to testing and assessment for READ-ACROSS BASED FILLING OF DEVELOPMENTAL TOXICITY DATA GAP FOR METHYL HEXANOIC ACID

Series on Testing and Assessment
No. 325

The corresponding annexes are available under the following cotes:
ENV/JM/MONO(2020)21/ANN2 andANN3.

JT03466608

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 325

CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR READ-ACROSS BASED FILLING OF DEVELOPMENTAL TOXICITY DATA GAP FOR METHYL HEXANOIC ACID

IOMC INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris 2020

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 37 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in twelve different series: **Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides; Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; Safety of Manufactured Nanomaterials;** and **Adverse Outcome Pathways**. More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/chemicalsafety/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organisations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

This publication is available electronically, at no charge.

Also published in the Series on testing and Assessment link

For this and many other Environment,
Health and Safety publications, consult the OECD's
World Wide Web site www.oecd.org/chemicalsafety/

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division
2, rue André-Pascal
75775 Paris cedex 16
France
Fax : (33-1) 44 30 61 80
E-mail : ehscont@oecd.org

Forward

OECD member countries have been making efforts to expand the use of alternative methods in assessing chemicals. The OECD has been developing guidance documents and tools for the use of alternative methods such as (Q)SAR, chemical categories and Adverse Outcome Pathways (AOPs) as a part of Integrated Approaches for Testing and Assessment (IATA). There is a need for the investigation of the practical applicability of these methods/tools for different aspects of regulatory decision-making, and to build upon case studies and assessment experience across jurisdictions.

The objective of the IATA Case Studies Project is to increase experience with the use of IATA by developing case studies, which constitute examples of predictions that are fit for regulatory use. The aim is to create common understanding of using novel methodologies and the generation of considerations/guidance stemming from these case studies.

This case study was developed by EU ToxRisk project (BIAC) for illustrating practical use of IATA and submitted to the 2019 review cycle of the IATA Case Studies Project. This case study was reviewed by the project team. The document was endorsed at the 4th meeting of the Working Party on Hazard Assessment in June 2020.

The following case study was also reviewed in the project in 2019:

1. CASE STUDY ON USE OF AN INTEGRATED APPROACH TO TESTING AND ASSESSMENT (IATA) AND NEW APPROACH METHODS TO INFORM A THEORETICAL READ-ACROSS FOR DERMAL EXPOSURE TO PROPYLPARABEN FROM COSMETICS, ENV/JM/MONO(2020)16.
2. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR SYSTEMIC TOXICITY ARISING FROM COSMETIC EXPOSURE TO CAFFEINE, ENV/JM/MONO(2020)17.
3. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR 90-DAY RAT ORAL REPEATED-DOSE TOXICITY OF CHLOROBENZENE-RELATED CHEMICALS, ENV/JM/MONO(2020)18.
4. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT TO INFORM READ-ACROSS OF P-ALKYLPHENOLS: REPEATED-DOSE TOXICITY, ENV/JM/MONO(2020)19.
5. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR PREDICTION OF A 90 DAY REPEATED DOSE TOXICITY STUDY (OECD 408) FOR 2-ETHYLBUTYRIC ACID USING A READ-ACROSS APPROACH FROM OTHER BRANCHED CARBOXYLIC ACIDS, ENV/JM/MONO(2020)20.
6. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR IDENTIFICATION AND CHARACTERISATION OF PARKINSONIAN HAZARD LIABILITY OF DEGUELIN BY AN AOP-BASED TESTING AND READ ACROSS APPROACH, ENV/JM/MONO(2020)22.

7. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR MITOCHONDRIAL COMPLEX-III-MEDIATED NEUROTOXICITY OF AZOXYSTROBIN - READ-ACROSS TO OTHER STROBILURINS, ENV/JM/MONO(2020)23.

These case studies are illustrative examples, and their publication as OECD monographs does not translate into direct acceptance of the methodologies for regulatory purposes across OECD countries. In addition, these cases studies should not be interpreted as official regulatory decisions made by the authoring member countries.

In addition, a considerations document summarising the learnings and lessons of the review experience of the case studies is published with the case studies:

REPORT ON CONSIDERATIONS FROM CASE STUDIES ON INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT (IATA) -Fifth Review Cycle (2019) -, ENV/JM/MONO(2020)24.

This document is published under the responsibility of the Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology.

Abstract / Synopsis / Executive summary

2-Methylhexanoic acid (MHA) is a compound for which developmental and reproductive toxicity test (DART) data is taken to be lacking. We have searched for structural analogues that have this data in order to explore the possibility to read across information of these source chemicals to MHA. The following structural related aliphatic carboxylic acids were selected that have *in vivo* developmental and/or reproductive toxicity data: 2-ethylhexanoic acid (EHA), 2-propylpentanoic acid (VPA), 2-propylheptanoic acid (PHA), 2-ethylbutanoic acid (EBA), 4-pentenoic acid (PA), 2-propyl-4-pentenoic acid (4-ene-VPA), and 2-dimethylpentanoic acid (DMPA). Some of these analogues proved to be clear developmental toxicants, i.e. VPA, PHA, EHA, and 4-ene-VPA, while others were identified as not being toxic to development, i.e. EBA, PA, and DMPA; i.e. they did or did not induce neural tube defects upon *in vivo* exposure. Thus, structural similarity alone doesn't allow a conclusion on the developmental toxicity of MHA. Therefore, we have also tested MHA and all the selected source chemicals in a battery of *in vitro* tests with clear relevance to developmental toxicity, i.e. the Zebrafish Embryo Test (ZET), mouse Embryonic Stem cell Test (mEST), iPSC-based neurodevelopmental model (UKN1), and a series of CALUX Reporter assays, that we combined with toxicokinetic models to calculate effective cellular concentrations and associated *in vivo* exposure doses. With these new approach methodologies (NAM) we wanted to explore whether they could correctly predict the *in vivo* developmental toxic properties of these aliphatic carboxylic acids, and thus could be used to predict the *in vivo* developmental toxicity of MHA itself. This data would also allow to further explore the relationship between structure and developmental toxicity within this series of aliphatic carboxylic acids. For that reason, we have also tested 2-methylpentanoic acid (MPA) in these NAM, despite the absence of *in vivo* data. We have also investigated the potential to inhibit histone deacetylase in ZET, mEST, and UKN1 models, as this enzyme is postulated to be the molecular initiating target leading to neural tube defects observed with these analogues.

The NAM results show that VPA, PHA, EHA, and 4-ene-VPA were correctly predicted as *in vivo* developmental toxicants, and EBA, and DMPA as non-developmental toxicants. The NAM results suggest that MHA may not be fully negative for developmental toxicity.

Table of Contents

Tables

Figures

1. Introduction

MHA is a compound for which DART data are lacking. This is considered a datagap of this compound if it is to be registered for market use within EU. We want to fill in this datagap via read across, hereby avoiding *in vivo* testing for DART, seen only as last resort by REACH (European Commission, 2006). Applying read across involves two major aspects, one of which is finding appropriate source chemicals, the other is selecting appropriate innovative *in vitro* models, and toxicokinetic and data analysis models for identifying and assessing the reproductive properties of these chemicals.

Finding appropriate source chemicals for MHA is in fact identifying (2-position branched) aliphatic carboxylic acids with adequate DART testing data. One such chemical is VPA (valproate, 2- propyl pentanoic acid), a human anticonvulsant drug, used in the treatment of among others epilepsy in humans. Human data indicate that the use of valproate during pregnancy is associated with foetal death and major congenital malformations. High dosages (>1000 mg/d) were associated with significant greater risks of congenital malformations than lower dosages. Malformations associated with maternal valproate use include neural tube defects (particularly spina bifida), cerebral, craniofacial, skeletal, cardiovascular and urogenital effects (hypospadias), developmental delays and cognitive deficits (Duncan *et al.*, 2001; Bowden, 2003; Koren, 2006; Kantola-Sorsa *et al.*, 2007; Thomas *et al.*, 2008; Ornoy, 2009; Tomson & Battino, 2009; Jentink *et al.*, 2010). Also in mice, at dose levels ranging from 200-600 mg/kg bw/day, neurodevelopmental effects (exencephaly, reduced head size, neural tube closure defects, microscopically disorganisation of neuroepithelium) were observed. At these concentrations a wide range of other developmental effects (among others foetal death, resorptions, growth retardations, craniofacial and skeletal malformations, cardiovascular effects, urogenital anomalies) as well as maternal toxicity (reduced food/water intake, reduced weights) were observed. For only one other structurally related aliphatic carboxylic acid such guideline animal data were found: EHA (Narotsky & Kavlock, 1994), showing a VPA-like profile of reproductive toxicity at 250-500 mg/kg bw in the rat. Such guideline studies could not be retrieved for other structurally related aliphatic carboxylic acids. However, for quite a number of such carboxylic acids specific prenatal developmental animal data were retrieved: this concerned a specific study protocol with mice consisting of monitoring induction of neural tube defects on gestation day 18, after a single dose exposure on day 8. Next to EHA, PHA, EBA, DMPA, PA, and 4-ene-VPA, also VPA was tested in this protocol, providing a rather unique set of uniform *in vivo* data for these structural analogues (see Costanza & Hartung, 2009). In this standardised mouse study VPA, EHA, PHA, and 4-ene-VPA were capable of inducing NTD; remarkably, the other analogues, i.e. EBA, DMPA, PA turned out to be negative in this respect. Thus, despite their common aliphatic carboxylic acid backbone, some appeared capable of inducing NTD, while others were incapable, using this specific animal model. A more detailed section on structural and phys-chem features of these aliphatic carboxylic acids is given in chapter 5 as part of the category hypothesis.

Selecting appropriate innovative *in vitro* models for identifying and assessing reproductive properties of chemicals is the other major aspect of this read across approach. Substantial efforts have already been undertaken to develop these methods for this regulatory endpoint (Adler *et al.*, 2011; Leist *et al.*, 2014). Clearly, a single assay cannot cover the whole mammalian reproductive cycle due to its inherent complexity, and, therefore, more recent studies have attempted to combine several *in vitro* assays into a test battery. Most notably,

the successive European projects ReProTect, and ChemScreen provided support for this approach [Schenk *et al.*, 2010; Piersma *et al.*, 2013), as did the US EPA's ToxCast program (Sipes *et al.*, 2011). Based on the approach developed in ChemScreen it could be demonstrated that the applied battery was able to successfully identify eleven out of twelve compounds with varying mechanisms of reproductive action. This battery approach included toxicokinetic modelling, proving a critical element in this assessment. Using the response data of the models used in ChemScreen we have selected a minimal essential battery for the current investigation, consisting of: Zebrafish Embryo Test (ZET), mouse Embryonic Stemcell Test (mEST; cardiac differentiation), iPSC-based neurodevelopmental model (UKN1), and a series of CALUX Reporter assays. The different models will be further described in chapter 5. , while the methods are detailed in Annexes to this report.

Appropriate innovative toxicokinetic models were developed for the purpose of transforming *in vitro* concentration response data into associated *in vivo* exposure doses. For achieving this, basically two models were needed: the first model should be capable of calculating free unbound *in vitro* target dose concentrations from nominal *in vitro* concentrations, while the second model should be capable of transforming this free unbound *in vitro* target dose concentrations into associated *in vivo* exposure doses. These *in vivo* exposure doses serve two purposes: one is to verify whether associated *in vivo* exposure doses are realistic exposures, and the other is to derive dose descriptors for risk assessment (e.g. to derive DNELs). The different models will be further described in chapter 7, while detailed descriptions are provided in Annexes to this report.

The induction of exencephaly, a neural tube defect, by aliphatic carboxylic acids like VPA is assumed to involve inhibition of histone deacetylase (HDAC) enzymes as an important possible initial step (Ornoy, 2009; Lloyd, 2013). There is clear evidence that this mode of action is also relevant to humans (Wiltse *et al.*, 2005). We have, therefore, also monitored HDAC inhibition by the aliphatic carboxylic acids investigated in this study in our experimental models to verify the correlation between HDAC inhibition and exencephaly induction.

The central question of this investigation was to explore the usefulness of NAM to substantiate and support the assessment of a chemical's developmental toxicity via read across.

2. Purpose

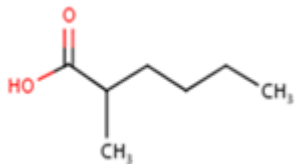
2.1. Purpose of use (targeted regulatory framework)

2-Methylhexanoic acid (MHA) is a compound which lacks information on developmental and reproductive toxicity test (DART) data. According to REACH this information is required for chemicals to be registered at tonnage levels exceeding 10 tonnes per annum. In order to fill this information gap the REACH Regulation encourages to apply where possible New Approach Methodologies (NAM), and to perform animal tests only as last resort. One of the options of fulfilling information requirements is via applying read across, that is by filling the data-gap of the chemical of interest by using test information of structurally related chemicals. The purpose of this report is to demonstrate that NAMs can effectively support the read across to filling a chemicals developmental toxicity data gap.

2.2. Target chemical / category definition / source chemicals

Target chemical

In this investigation 2-methyl hexanoic acid (MHA) is being investigated. Used synonyms are 2-methylcaproic acid, 2-hexanecarboxylic acid, and α- methylcaproic acid. The molecular formula is $C_7H_{14}O_2$, and its structure formula is as follows:



The structure descriptors for MHA are:

Smiles code: CCCCC(C)C(=O)O

InChi: S/C7H14O2/c1-3-4-5-6(2)7(8)9/h6H,3-5H2,1-2H3,(H,8,9)

InChiKey: CVKMFSAVYPAZTQ-UHFFFAOYSA-N

Chemical identifiers of MHA are:

CAS. No: 4536-23-6

EINECS: 224-883-9

TSCA INV: 22160-12-9

FDA UNII: LUK37N0QM8

MHA is a 2-branched aliphatic carboxylic acid, whose physical-chemical characteristics (molecular weight, melting and boiling point, logPoW, water solubility, pKa, vapour pressure, and Henry's Law constant) are: (all also described in the data-matrix)

Molecular weight: 130,1

Melting point: 26,62

Boiling point: 215,45

LogPoW: 1,8 (exp.); 2,47 (pred.)

Water solubility: unknown / expected around 2000 mg/L (for EHA, 25 ºC; see Annex I)

pKa: 4,8 (pred.)

Vapour pressure: unknown / expected within 0.18-0.03 mg/L mm Hg (25 ºC; see Annex I)

Henry's Law constant: 0,229 (pred.)

Category definition

It is thought that the health effects data gap of MHA could best be filled by reading across from chemicals having a similar structure and that have this required health effects properties defined. Most suitable, therefore, we consider other 2-branched aliphatic carboxylic acids, having comparable chain lengths. This group description should also define comparable phys-chem property ranges that could be considered acceptable to accept these chemicals as analogues in this category to allow and justify the performance of a read across within this group.

Source chemicals

The source chemicals, and positive and negative controls (named 'outliers' in the data-matrix) are as follows: source chemicals are ethyl butyric acid (EBA), ethyl hexanoic acid (EHA), propyl pentanoic acid (VPA; Valproic acid), and propyl heptanoic acid (PHA), while controls ('outliers') are pentenoic acid (PA), dimethyl pentanoic acid (DMPA), and propyl pentenoic acid (4-ene-VPA); see below Table 1 with CAS.NRs and *in vivo* responses. Methyl pentanoic acid (MPA) was included for its structural similarity, though it lacked *in vivo* data. Chemical structures and more chemical identifier characteristics are provided in the data-matrix in Annex I.

2.3. Endpoint(s) for which the read-across is performed

MHA lacks data on developmental toxicity. In this investigation this data gap is to be filled, i.e. the developmental properties of MHA are to be identified and characterised via a read across assessment, using structurally similar chemicals that have this developmental properties information.

2.4. Exposure information

Not applicable for the read-across case described here.

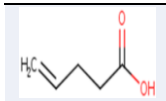
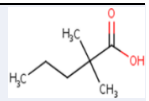
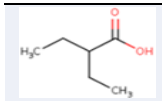
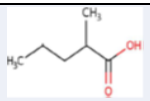
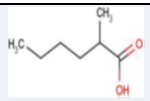
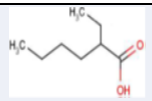
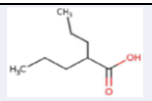
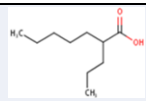
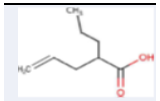
3. Hypothesis for the category approach and selected source chemicals

MHA is a 2-branched aliphatic carboxylic acid that lacks developmental toxicity data. In the above chapter, it is indicated that we consider other 2-branched aliphatic carboxylic acids with comparable chain length suitable source chemicals for reading across these health effects to MHA, provided that these chemicals themselves have adequate developmental toxicity data. It is assumed that by keeping as strict inclusion criteria for source chemicals only aliphatic carboxylic acids branched at position 2, and by selecting aliphatic carboxylic acids with comparable chain length, that both phys-chem properties, that drive toxicokinetic processes, and structural features, that may trigger toxicodynamic processes, the associated toxicokinetic and -dynamic properties for these source chemicals will be only minimally deviating.

It is recognised that the category only contains one member with negative *in vivo* data; which introduces an uncertainty. Broadening the category criteria, however, to potentially increase the number of negative (and positive) members was considered, but not decided for, as this would certainly increase the uncertainty of the structural similarity basis of the read across.

The source chemicals selected in this study allow the read across to MHA to be classified as a mere interpolation rather than extrapolation, as MHA appears to be a carboxylic acid with intermediate aliphatic chain lengths: see below table depicting target and source chemicals. The target chemical MHA has two aliphatic chains at position 2 with carbon chain lengths of resp. n=1, and n=4: see Table 1.

Table 1. Target and source chemicals of this read across study depicted with their chain lengths; also some structurally closely related control chemicals are depicted.

CAS-RN Abbreviated name	591-80-0 PA	1185-39-3 DMPA	88-09-5 EBA	97-61-0 MPA	4536-23-6 MHA	149-57-5 EHA	99-66-1 VPA	31080-39-4 PHA	1575-72-0 4-ene-VPA
Structure									
Chain lengths at position 2	3 / 0 / 0	3 / 1 / 1	2 / 2 / 0	3 / 1 / 0	4 / 1 / 0	4 / 2 / 0	3 / 3 / 0	5 / 3 / 0	3 / 3 / 0
Identity / function	control	control	source	*in vitro*	target	source	source	source	control
In vivo response / potency	neg	neg	neg	no data	?	pos: +	pos: +++	pos: +++	pos: ++
Structural similarity RDKit fingerprint (%)	-	-	59	83	100	82	69	77	-

From the above table one can see that source chemicals directly on the left (MPA) and right hand side (EHA) of MHA are having only one carbon-atom chain length difference, and those more distant to MHA some may differ two carbons in length in one of the chains. The structural similarity of source chemicals to the target is provided as well, using RDKit fingerprints (Daylight-like topological fingerprint, and Tanimoto similarity metric) showing the closest similarity of MPA and EHA relative to MHA. In de data-matrix also results of some other similarity tools (FCFP4 Fingerprint [Circular fingerprint based on the Morgan algorithm and feature invariants], Layered Fingerprint [An experimental substructure-matching fingerprint], and RDKit Descriptors are shown: all showing the closest similarity of MPA and EHA. The control chemicals have less structural similarity to the target; their role in this read across assessment is to provide additional positive and negative *in vitro* effect profiles in the applied models.

When this composition of the category is compared to the scenario options in the Read Across Assessment Framework (RAAF; ECHA, 2017), scenario number 4 seems most applicable here: a category approach, where the read across is based on different compounds that have the same type of effect(s), and where quantitative variations may exist in the strength of effect(s) observed among source substances. The RAAF indicates that prediction for the target chemical here could be based on an observed regular pattern within the category or on a worst-case approach.

Table 1 also shows the *in vivo* response data of these chemicals: some appear negative *in vivo* ('neg'), some are positive ('pos'), though with different potency. This *in vivo* data is further described below in section 3.3.

An interesting observation is that the analogues on the right side of MHA are all positive *in vivo*, while of the two analogues on the left, the one with *in vivo* data, i.e. EBA, is negative. As indicated above we have also included 'outlier' structures because, though structurally less similar, they have relevant *in vivo* data. Because of the *in vivo* data of the category analogues, we have positioned the *in vivo* negative control chemicals PA, and DMPA on the left, and the *in vivo* positive control 4-ene-VPA on the right side of this category list.

3.1. Chemical identity and composition

The chemical identity of the target chemical is described in section 'Target chemical(s) / category definition' in chapter 2.2. The chemical identity of the source and control chemicals are described in the data-matrix (folders 'Source and Target cmpds', and 'Physchem'). All chemicals investigated in this report are of analytical grade and purchased by one partner as central facility (IFADO, Dortmund, Germany), and from there distributed to the different collaborating laboratories.

3.2. Physical-chemical properties and other molecular descriptors

The physico-chemical properties and other molecular descriptors of MHA and the source and control chemicals are provided in Table 1b (and listed in the data-matrix in Annex I in folder 'Physchem'). From this it appears that from fully left (EBA) to fully right (PHA) within the category there is a trend for all these parameters (rounded values): molecular weights of the sources range from 116 to 172 g/mol, melting and boiling points increase from 15 to 59, and from 196 to 269 ºC, respectively. LogPow ranges from 1,7 to 3,2 (experimental) or from 2,0 to 3,9 (predicted). Water solubility shows, for only some available, a decrease going from left (EBA) to right (PHA), i.e. from 18 to 0,3 g/L. The pKa constant values appear rather similar: experimentally derived values go from 4,7 (EBA) to 4.6 (VPA), while predicted values are similar for all 6 members in the category, i.e. 4.8, indicating these are all weak acids. Vapour pressures reduce when going from left to right (0.19 for EBA to 0,005 mmHg for PHA at room temperature. Henry's Law constant was experimentally only determined for EHA, being 0,29 Pa m^3/mol, while predicted values increase from 0,16 for EBA to 0,54 Pa m^3/mol for PHA. The conclusion is that these compounds are all non-volatile, well to reasonable water soluble, will not bio-accumulate, and are weak acids of quite comparable strength.

Table 1b. Physchem properties of target and source chemicals of this read across study

CAS-RN Abbreviated name	591-80-0 PA	1185-39-3 DMPA	88-09-5 EBA	97-61-0 MPA	4536-23-6 MHA	149-57-5 EHA	99-66-1 VPA	31080-39-4 PHA	1575-72-0 4-ene-VPA
Molecular weight	102,133	130,187	116,08	116,08	130,1	144,12	144,12	172,15	142,2
Melting point (°C; pred. EPISUITE)	14,76	24,61	15,24	15,24	26,62	37,72	37,72	59,15	36,44
Boiling point (°C; pred. EPISUITE)	187,75	207,77	195,8	195,8	215,45	234,2	234,2	268,98	232,83
logPow (exp. EPISUITE)	1,39	-	1,68	1,8	1,8	2,64	2,75	3,2	-
logPow (pred. EPISUITE)	1,56	2,43	1,98	1,98	2,47	2,96	2,96	3,94	2,82
Water solubility (mg/L; exp.)	24000 (25°C)	-	18000 (20°C)	-	-	2000 (20°C)	2000 (20°C)	275.6 (25°C)	-
pKa (exp.)	-	-	4.71	-	-	4.7	4.6	-	-
pKa (pred. ACD/Percepta)	4,84	4,9	4,8	4,8	4,8	4,8	4,8	4,8	4,7
Vapour pressure (mm Hg; exp.)	-	-	0.188 (25°C)	-	-	0.03	0.0458 (25°C)	0.0048	-
Henry's Law Constant (Pa m³/mol; pred. EPISUITE (bond method))	0,130	0,229	0,162	0,172	0,229	0,304	0,304	0,535	0,226
Henry's Law Constant (Pa m³/mol; exp. EPISUITE)	-	-	-	-	-	0,289	-	-	-

experimental (exp.) or predicted (pred.) values and source models listed; '-': no data

3.3. *In vivo* data source chemicals

As shown in Table 1 for 4 of the 5 source chemicals, and for the controls (named 'outliers' in report template) we do have *in vivo* data on developmental toxicity. For MPA we do not have this data but we included it in this study to explore the relationship between structure and developmental toxicity within this series of aliphatic carboxylic acids, as it may show support for a trend that may come out of this investigation.

The *in vivo* data of this 4 source and 3 control chemicals all come from a standard teratogenicity protocol developed by Nau and colleagues to further study the induction of neural tube defects by VPA: this is the most apparent lesion in mouse, and also observed in humans exposed to VPA (Nau *et al.*, 1981, 1986). In short, in this standard teratogenicity protocol female NMRI mice (28-32 g), housed under specific-pathogen-free conditions (diet and tapwater *ad libitum*), were mated with males for periods of 2 hr (8 to 10 AM), and a subsequent 24 hr examined for vaginal plugs; if found, this was designated as Day 0 of gestation. Groups of 15 pregnant animals were exposed to test compounds at day 8 of gestation (by single sc injection) at 400 or 600 mg/kg bw in water (10 ml/kg bw), and were examined on day 18 of gestation for implantation sites, and each live foetus was individually weighed and inspected for the presence of the neural tube defect exencephaly. Effects on visceral and skeletal systems were not examined. The outlier substances DMPA, 4-ene-VPA, and PA were tested in this same teratogenicity protocol (Nau *et al.*, 1981, 1986, 1991; Eikel *et al.*, 2006; Courage-Maguire *et al.*, 1997; Dearden *et al.*, 2013). These studies showed that aliphatic carboxylic acids have different potencies of inducing neural tube

defects, and that some of them even were inactive in this respect. Table 2 shows the potency of induction of exencephaly by the aliphatic carboxylic acids: of the source chemicals, EBA is negative, while EHA, VPA and PHA are positive, though at different potencies, as indicated below.

Table 2. Teratogenic potency grading in NMRI mouse exencephaly model (Eikel *et al.*, 2006)

Teratogenic potency	Dose range (mmol/kg bw)	Exencephaly rate (%)	Description
0	> 3.0	0	No teratogenic potency detectable
+	2.0 - 3.0	1-5	Low teratogenic potency
++	2.0 - 3.0	5 - 25	Lower teratogenic potency than VPA
+++	2.0 - 3.0	25 - 60	Equal teratogenic potency to VPA
++++	1.0 - 2.0	40 - 60	Higher teratogenic potency than VPA

The above NMRI mouse study results are included in the data-matrix in Annex I under the header 'target endpoint 2'.

By comparing structure and potency in Table 1, the data suggest that side-chain length is related to potency: PHA, having the longest side chains within the category has the highest potency, while EBA with the smallest side chains is inactive. EHA, with in between side chain length has the lowest potency. For MHA, and MPA, that are in between EHA, and EBA with regard to chain length, we have no *in vivo* data, and it is, therefore, unclear whether they possess any potential for developmental toxicity *in vivo*: this read across case study is to fill this gap. There is no clear explanation for this structure-related potency, other than that chain-length probably is a factor in affinity for binding to the critical target that leads to the *in vivo* observed effects.

Only for VPA and EHA other *in vivo* developmental toxicity study data were available. For VPA, in fact there is a wealth of data available. A series of developmental toxicity studies in mice in which neurodevelopmental parameters were included are described (Paulson *et al.*, 1985; Padmanabhan, 1996; Sonoda *et al.*, 1990; Turner *et al.*, 1990). In these studies, at dose levels ranging from 200-600 mg/kg bw/day, neurodevelopmental effects (exencephaly, reduced head size, neural tube closure defects, microscopically disorganisation of neuroepithelium) were observed. The lowest LOAEL that was described in mice was 200 mg/kg bw/day (Padmanabhan *et al.*, 1996). The lowest NOAEL that was described in mice was <200 mg/kg bw/day. At this concentration, exencephaly and neural tube closure effects were observed. In addition, in the study of Paulson *et al.* (1985), and Sonoda *et al.* (1990), exencephaly was observed at 340 and 560 mg/kg bw/day and 600 mg/kg bw/day, respectively. In general, at concentrations that induced neurodevelopmental effects, a wide range of other developmental effects (among others foetal death, resorptions, growth retardations, craniofacial and skeletal malformations, cardiovascular effects, urogenital anomalies) as well as maternal toxicity (reduced food/water intake, reduced weights) were observed. In addition to the neurodevelopmental toxicity studies, a series of developmental toxicity studies was described in which no neurodevelopmental parameters were studied (ECHA website). Overall, the lowest LOAEL for 'general' developmental toxic effects in mice was 160 mg/kg bw/day.

There were two rat studies in which neurodevelopment parameters of VPA were studied (Ong *et al.*, 1983; Frisch *et al.*, 2009). In the study of Ong *et al.*, no neurodevelopmental effects were observed at concentrations as high as 600 mg/kg bw/day whereas in the study of Frisch *et al.*, functional and structural effects were observed at 720 mg/kg bw/day, which is considered as the lowest LOAEL for neurodevelopmental effects of VPA in rats. In

addition, a series of developmental toxicity studies was performed in which no neurodevelopmental parameters were included. Overall, general developmental effects of VPA in rats were observed at 100 mg/kg bw/day and higher.

For VPA one rabbit study was reported that investigated neurodevelopment parameters. In this study, the NOAEL for neurodevelopmental toxicity was >350 mg/kg bw/day, whereas the NOAEL for developmental toxicity was 50 mg/kg bw/day (Petrere *et al.*, 1986). No maternal toxicity was reported. In a prenatal developmental toxicity study with New Zealand White rabbits (without neurodevelopmental endpoints), the NOAEL for prenatal developmental toxicity was 150 mg/kg bw/day (ECHA website). At higher concentrations of 250 and 350 mg/kg bw/day, the incidence of resorptions was increased and foetal weight was decreased.

Hendrickx *et al.* (1988) reported a study with Rhesus monkeys treated with sodium valproate during gestation days 21-50. The NOAEL for neurodevelopmental toxicity was 75 mg/kg bw/day, at >100 mg/kg bw/day, the incidence of foetuses with a reduced head circumference was increased. Other developmental toxic effects were observed at concentrations higher than 20 mg/kg bw/day (craniofacial skeletal effects, growth retardation, embryo- and foetal mortality). At all concentration tested, maternal toxicity was observed.

For EHA, a series of rat pre-natal developmental toxicity studies and a series of rat reproductive toxicity studies, including an extended-one generation reproductive toxicity study (EOGRTS) according to OECD guideline 443, was available (ECHA website). In one of the pre-natal developmental toxicity studies, dilation of the lateral ventricles of the brain was observed in foetuses of dams treated with 500 mg/kg bw/day, which might be considered as a neurodevelopmental toxic effect. At this concentration, other developmental effects (reduced foetal weight, skeletal effects) as well as maternal toxic effects were observed. In the EOGRTS study, no neurodevelopmental effects were observed at concentrations up to and including 800 mg/kg bw/day (ECHA website). At this concentration, no other reproductive- and developmental toxic effects were observed whereas slight maternal toxicity (body weight, food consumption, kidney and liver pathology) was observed. In two other reproductive- and developmental toxicity studies neurodevelopmental parameters were not included but developmental toxicity (effects on pup weights, litter size, physical development) was observed at dose levels as low as 100 mg/kg bw/day (ECHA website; Pennanen *et al.*, 1992; Bui *et al.*, 1998).

In a prenatal developmental toxicity study with New Zealand White rabbits, the animals were treated with 0-250 mg/kg bw/day of EHA (ECHA website). No developmental toxic effects were observed, whereas maternal toxicity was observed at dose levels of 25 mg/kg bw/day and higher.

Relevant NOAELs of the above studies are included in the data-matrix in Annex I under the header ‘target endpoint 1’.

Only for VPA also human data is available indicating that the use of valproate during pregnancy is associated with foetal death and major congenital malformations. High dosages (>1000 mg/d) were associated with significant greater risks of congenital malformations than lower dosages (Diav-Citrin *et al.*, 2008; Jentink *et al.*, 2010). Malformations associated with maternal valproate use include among others neural tube defects (particularly spina bifida), cerebral, craniofacial, skeletal, cardiovascular and urogenital effects (hypospadias). The mouse model appears a suitable animal model reflecting similar findings such as among others neurodevelopmental effects (exencephaly,

reduced head size, neural tube closure defects) at dose levels ranging from 200-600 mg/kg bw/day (Nau *et al.*, 1991; Turner *et al.*, 1990; Sonoda *et al.*, 1990).

As a final remark to this study data, it is noted that some of the aliphatic carboxylic acids have asymmetric carbon atoms, and thus may exist as racemates: i.e. MPA, MHA, EHA, PHA, and 4-ene-VPA. The *in vivo* data in the NMRI mouse model were generated with the racemates, as far as data on this was retrievable; therefore, all *in vitro* tests described in this report have been performed with racemates of these chemicals as well.

3.4. Kinetics: Absorption, distribution, metabolism and excretion

In the folder 'ADME-Toxicokinetics' in the data-matrix in Annex I the ADME characteristics 'unbound fraction' (fu), hepatic intrinsic and *in vivo* clearance (CL_{int} and CL, respectively), as well as total and unbound steady state volume of distribution (Vss and Vu,ss, respectively) for the category members (target and source chemicals) and controls (named 'outliers' in report template) are listed. The models from which these values were derived will be described below.

Physiologically based pharmacokinetic modelling

Physiologically-based pharmacokinetic (PBPK) modelling and simulation was applied as part of the read-across with two primary aims:

1) To predict the pharmacokinetics of source and target compounds in the read-across based on *in vitro* to *in vivo* extrapolation (IVIVE) PBPK and allometric scaling.
2) Determine the concentrations of compounds in target tissues, namely here the foetal-placental compartment, and through reverse dosimetry translate *in vitro* toxicity assay data to an oral equivalent dose.

PBPK models for read-across compounds were developed in the human and animal Simcyp Simulators (V17r1, Certara UK Ltd. Simcyp Division, Sheffield, UK; www.simcyp.com) production of each version of the Simcyp simulator has been described in detail (Jamei *et al.*, 2013).

In developing these PBPK models for the read-across study, the following aspects were considered as suggested in the WHO PBPK guidance (WHO publication Harmonisation Project Document No. 9. Characterisation and Application of Physiologically based PharmacoKinetic Models in Risk Assessment):

1) The source or target compound was assumed to be the toxic moiety (i.e. plasma and tissue levels of formed metabolites were not routinely considered in the PBPK models).
2) Metabolism is thought to be the major clearance pathway of the compounds in this read-across study and the metabolic clearance in humans was predicted using an IVIVE approach, scaling *in vitro* intrinsic clearance determined in human primary hepatocytes.
3) The physiology (i.e. tissue weights and blood flow rates) of the species of interest were the default values in the Simcyp human and animal simulators (V17r1, Certara UK Ltd. Simcyp Division, Sheffield, UK).

Development of PBPK models

Where sufficient input data was available, PBPK models were developed for each of the source and target compounds; modelled compounds are EHA, MPA, MHA, EHA, and VPA. The PBPK model development work-flow is summarised in Figure 1 below and exemplar model input parameters are shown in Table 3; the parameters for each compound specific PBPK model in each of the species of interest are tabulated in Annex III.

Figure 1. Human and model species physiologically-based pharmacokinetic model development work-flow.

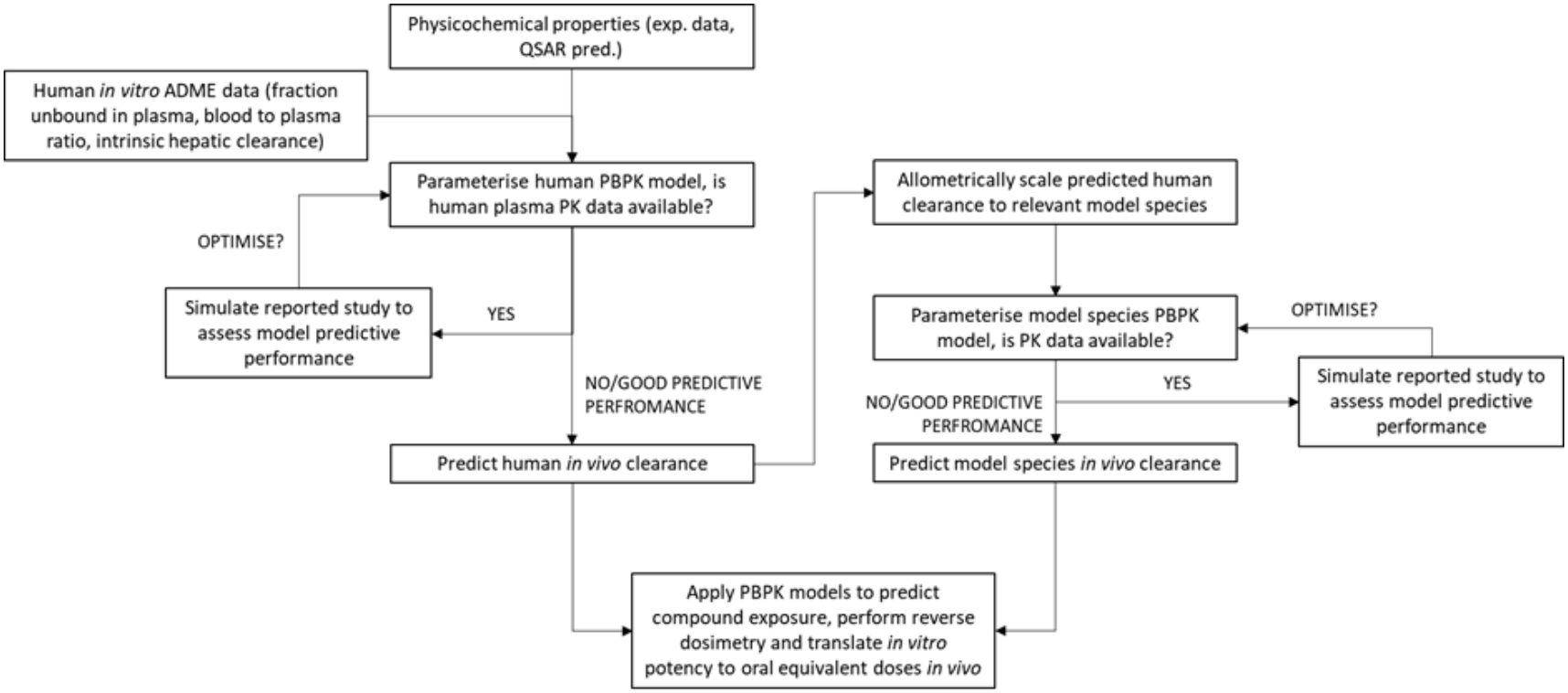


Table 3. Input parameters for development of human valproic acid PBPK model.

Parameter	Value	Method	Source/Reference
MW [g/mol]	144.21		EPI-Suite (v4.1, US-EPA)
logP	2.75	exp.	EPI-Suite (v4.1, US-EPA)
Compound Type	Mono-protic acid		
pKa	4.8		ACD/Percepta (2012 release, build 2254, ACD Labs)
TPSA (Å^2)	37.3		https://pubchem.ncbi.nlm.nih.gov/compound/3121#section=Chemical-and-Physical-Properties
Hydrogen bond donors	1		https://pubchem.ncbi.nlm.nih.gov/compound/3121#section=Chemical-and-Physical-Properties
fu	0.138	Pred. (min)	Lhasa
	0.310	Pred. (max)	CORAL model
B/P ratio	0.55	assumed	
fa	0.996	Pred. (Simcyp v17r1)	(Winiwarter *et al.*, 2003; Winiwarter *et al.*, 1998)
ka(h^{-1})	2.546	Pred.(Simcyp V17r1)	
fu_{Gut}	1	assumed	
CL_{int} (µl/min/10^6 cells)	0.219	Exp.	HREL co-culture system; Cyprotex data (CYP1440)
Hepatocyte binding (fu heps)	0.954	Pred.	(Kilford *et al.*, 2008)

Legend: fu = fraction unbound in plasma; CL_{int} = intrinsic clearance rate in primary hepatocytes; ka=absorption rate; fa= fraction absorbed; B/P ratio= blood plasma ratio.

The following general assumptions were made in the development of PBPK models for this read-across:

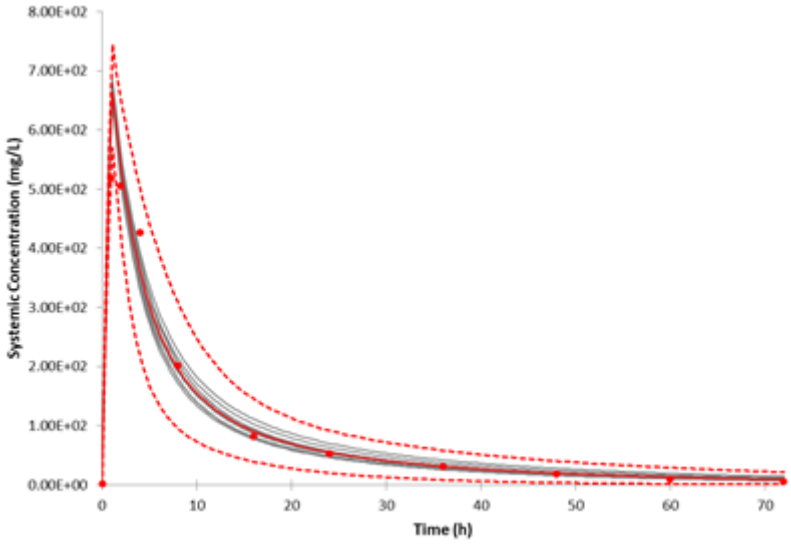
1) The metabolic clearance of target and source compounds is assumed to be linear (i.e. not to saturate) over the range of doses simulated.
2) Plasma protein binding is assumed to be linear (i.e. not to saturate) over the range of doses simulated
3) There is no intestinal metabolism or metabolism in any of the tissues in the PBPK model apart from in the liver.
4) Potential cleavage of the compounds in the blood was not considered in the models.

The model input parameters intrinsic clearance in human hepatocytes (CL_{int}), the fraction unbound in plasma (fu), and blood to plasma ratio (BP) were determined experimentally *in vitro*. While experimental data for CL_{int} was successfully generated for each of the read-across compounds, the determination of fu and B/P was confounded by analytical issues. For fu this data gap was resolved through the use of QSAR models to predict PBPK model input parameters. Acidic compounds, significantly ionised at physiological pH, do not extensively distribute into erythrocytes, and given their physicochemical properties, the compounds in this study are all significantly ionised at the pH of plasma. It was therefore assumed that B/P = 0.55 for all case study compounds, representing minimal distribution into erythrocytes.

In this case study, VPA was the only source compound with sufficient human pharmacokinetic data available to verify predictive performance of the compound specific PBPK model. The input parameters for the human VPA PBPK model are summarised in Table 3, simulations with this model of Valproic acid following dosing at 30mg/kg show good recovery of the published clinical data (Figure 2). PBPK model performance verification for VPA was used to inform the development of PBPK models for other source

and target compounds, and assess the validity of modelling approach adopted across the category.

Figure 2. PBPK modelling for human plasma concentration of VPA following 30mg/kg IV dose



100 virtual individuals (proportion of females = 0.1) were simulated and randomly assigned to 10 trials; mean of individual trials (grey line), mean of simulated population (red line), percentile (5th, 95th, red dashed lines), observed experimental data for VPA (red points; (Georgoff *et al.*, 2018).

Sensitivity analyses, parameter uncertainty and model assumptions

There was uncertainty around some model input parameter values, or model assumptions. For these parameters, sensitivity analysis was conducted, using Valproic acid as an exemplar compound. The impact on the simulation results, namely systemic plasma exposure (plasma AUC) was evaluated for the parameters CL_{int}, fu, B/P, cardiac output, hepatic blood flow, and renal blood flow. Sensitivity analysis results are presented in the Table 4. Uncertainty is a subjective assessment of how reliable the input parameters are. A formal uncertainty analysis as suggested in the WHO PBPK guidance is difficult to perform with a bottom up PBPK model as the ratio of median to 95th percentile reflects a measure of variability rather than true uncertainty as could be obtained if the PBPK model parameters were fitted to an observed dataset.

Table 4. Summary of model input parameter sensitivity analysis

		High	Medium	Low
SENSITIVITY	High	plasma fu		
	Medium			
	Low		*in vitro* CL_{int}	B/P

a sensitivity ratio of 1 implies that a 1% change in the value of an input parameter leads to a 1% change in dose metric prediction. High (ratio greater than or equal to 0.5), medium (ratio greater than or equal to 0.2 but less than 0.5) or low (ratio greater than or equal to 0.1 but less than 0.2); parameters with sensitivities less than 0.1 are not tabulated.

Sensitivity analysis showed only low sensitivity of the model to variability in B/P and simulations with the VPA model using the assumption that the B/P = 0.55 showed good recovery of the observed data (Figure 2).

As experimental fu input data were not available for the source (excluding VPA) and target compounds QSAR predictions were used. Reported values for the plasma protein binding of VPA vary as a result of concentration dependent (saturable) binding (Ogungbenro *et al.*, 2014). Sensitivity analysis covering the range of QSAR predicted plasma protein binding values for VPA (fu = 0.1-0.4), shows that predicted total plasma exposure (AUC) decreases with increasing fu parameter values. This is an expected behaviour of the model, as an increasing fraction of compound is unbound, a higher fraction of compound is free for metabolic clearance. While the non-linearity in plasma protein binding for VPA is described in the literature, there is insufficient data to inform this modelling decision across all the read-across compounds. In addition, the assumption of saturable, non-linear protein binding could result in an under-prediction of exposure with increasing dose. To capture a 'worst-case' predicted systemic exposure, we decided to model each compound at the lowest and highest predicted fu value assuming linear, non-saturable plasma protein binding.

The *in vitro* intrinsic hepatic clearance (CL_{int}) value is considered medium uncertainty based on analytical challenges of measuring this parameter *in vitro* for these compounds. The model showed no significant sensitivity to the other parameters tested.

Reverse dosimetry to establish oral equivalent dose in human and mouse

Biokinetic modelling (see Annex III) was used to predict the intracellular concentrations corresponding to effective concentrations determined experimentally based on nominal treatment concentrations used in *in vitro* assays. For assay systems conforming to steady-state assumptions, a biokinetic model accounting for cellular composition and physicochemical properties was used to predict intracellular concentration. Alternatively, a simplified approach was used to predict free-concentrations in treatment medium (Annex III); predictions reported in Annex I. For assays performed in zebrafish embryo, a predictive model was used to predict total internal embryo concentrations.

Oral equivalent doses (OED; mg/kg) were subsequently predicted for each *in vitro* endpoint in human and mouse using the PBPK models developed. As described above, the minimum and maximum predicted values of fu were used, resulting in a range of equivalent doses for each endpoint, in each species. OED was defined as the oral dose resulting in a peak unbound plasma, or tissue concentration equal to the corrected or nominal EC_{10}/LEC (μM) determined *in vitro*.

Metabolism

Information on metabolism is available for only two of the six investigated category members: i.e. VPA, and EHA. Meteor Nexus, a knowledge based system, was used to predict the biotransformations for the remaining 4 members. The detailing report is included as Annex III.

VPA metabolism *in vivo* as well as *in vitro* is extensively studied, well documented and reviewed for a number of animal species and human (Silva *et al.*,2008; Ghodke-Puranik *et al.*,2013). Metabolites can be categorised as those arising from CYP-catalysed oxidation at alkyl side chains (microsomes, endoplasmic reticulum), those arising from beta-oxidation pathways (mitochondria) and those occurring as a consequence of phase II reactions of

conjugation (endoplasmic reticulum and cytosol). VPA is highly protein bound (87–95%) resulting in low clearance (6–20 ml/h/kg). Its major urinary metabolite is the valproate glucuronide accounting for 30-50% of an administered dose. beta-Oxidation is the most important oxidative biotransformation type (>40%) for VPA with CYP-based hydroxylation/dehydrogenation (15-20%) playing a secondary role. Some 50-70 different metabolites have been suggested in the literature for VPA.

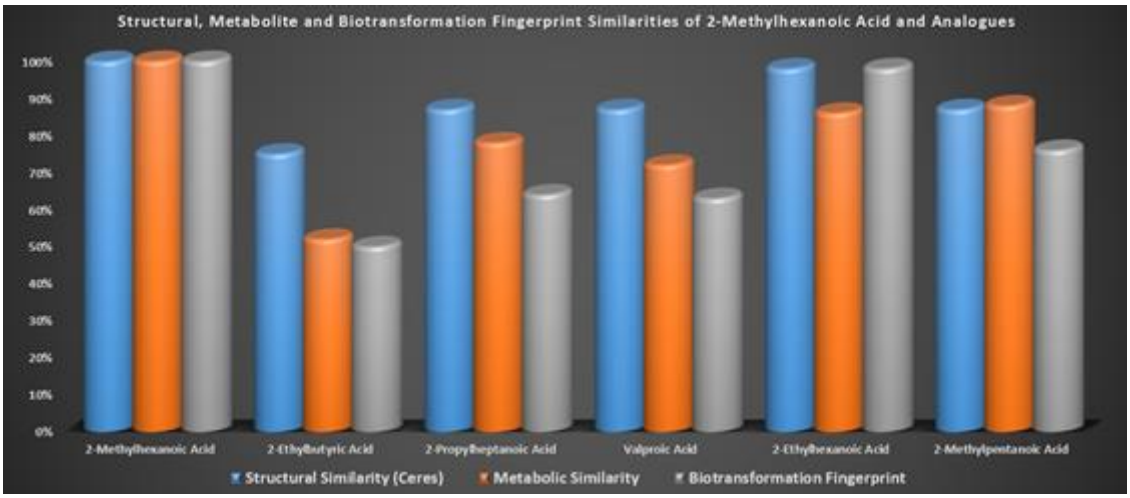
EHA (also a metabolite of the plasticiser bis(2-ethylhexyl)phthalate) is reasonably well studied (Pennanen *et al.*, 1991, 1996; English *et al.*,1998; Walker and Mills, 2001). Data is available from both *in vitro* and *in vivo* experiments in human, rat, rabbit, mouse, monkey, guinea pig and dog. Like VPA, metabolism is by glucuronide conjugation, beta-oxidation and CYP-mediated hydroxylation/dehydrogenation.

VPA has an extensive and complex metabolic profile. Meteor Nexus is able to predict all Phase II bio-transformations with the exception of the adenylate (adenosine monophosphate) conjugate, that is reported in a couple of studies, though not at significant levels. Whilst beta-oxidation is predicted for most analogues, the assigned score is always zero. The reason for this is that the Meteor choice of nearest neighbour algorithm requires the metabolite structure in the data set, in the case of beta-oxidation (and in respect of the meteor definition of this biotransformation) the final product that results after the final thiolytic cleavage and CoA deconjugation; this final thiolytic cleavage step has been hinted at in the literature but never observed for valproic acid or any of its analogues. The theoretical structures and intermediates are available for inspection in the Meteor Nexus tree. For substrates with a branched propyl chain, Meteor Nexus predicts conjugation with glutathione (biotransformation 506). Within this pathway, there are intermediates which arise as a result of sequential Cyp-oxidation (terminal dehydrogenation) and beta-oxidation (internal dehydrogenation), which is followed by 1,6-conjugate addition of glutathione to the 2-propyl-penta-2,4-dienoic acid. This reaction has been observed for valproic, and 4-ene-VPA but would in any case seem to be peculiar to propyl groups as a terminal dehydrogenation is involved. All of the major Cyp oxidations seen in the literature are predicted by Meteor Nexus, including some of the non-obvious metabolites such as those resulting from lactone formation. However, it is necessary to examine the larger, multi-generational trees in order to observe the predictions of these downstream metabolites. Understanding the limitation of the zero score for beta-oxidations, Meteor predicts well for this class in respect of the major and most often observed metabolites.

The general (and expected) trend in prediction (and limited observation) is that the relative amount of carbo-aliphatic oxidation increases as the length of the alkyl chain and the side branch increases. Some degree of conjugation (by a combination of glucuronic acid, glycine, glutamine and taurine) is always predicted and these are usually the most significant contributors to clearance of these compounds *in vivo*.

Below the structural and predicted metabolic (concordant metabolite) and biotransformation fingerprint similarities for target and category members are depicted in Figure 3. Concordant metabolites are defined here as those arising via the same biotransformation occurring at comparable positions with the target analogue. The analysis of physico-chemical and structural properties indicates that all compounds are covered in the domain of the Meteor Nexus knowledge base.

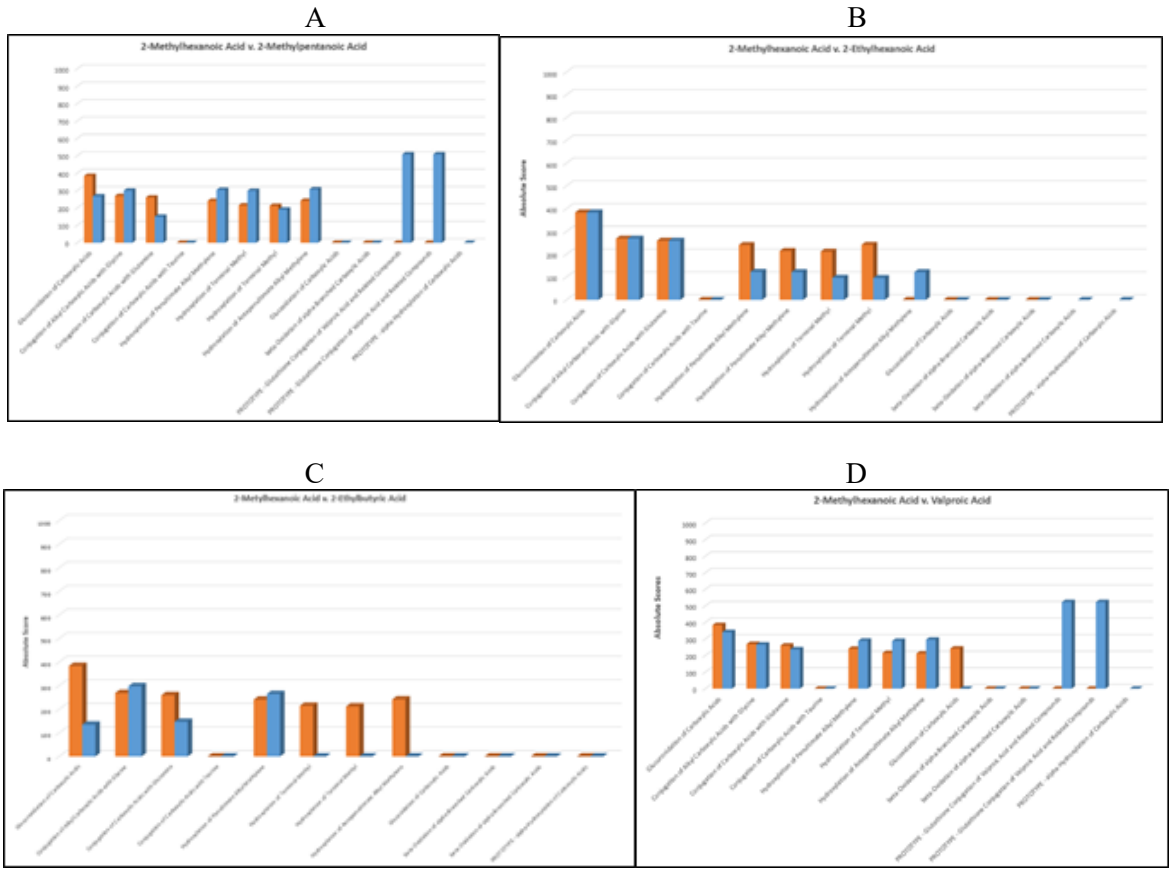
Figure 3. Comparison of Structural, Metabolic (Concordant Metabolite) and Biotransformation Fingerprint Similarities for Target and Analogue Compounds



A more detailed metabolism pattern (first generation) between the target MHA and the two nearby category members, MPA, EHA, VPA, and EBA is depicted below (Figure 4).

Figure 4. Predicted metabolism for target and source chemicals in the category by knowledge based system Meteor Nexus

MHA vs MPA (A), MHA vs EHA (B), and MHA vs EBA (C), and MHA vs VPA (D): orange bars: target chemical; blue bars: source chemical (depicted absolute scores are not normalised).



From this data, it is concluded that regarding similarity on biotransformation fingerprint (BF), and concordant metabolite (CM), MHA is most similar to EHA, and MPA, with BF and CM scores of 0.86 and 0.98, and 0.88 and 0.76, respectively. Similarity to VPA and EBA was less, with BF and CM scores of 0.63 and 0.72, and 0.50 and 0.52, respectively. More details on these similarities between MHA and the other category members, and more details on methods, and data and method reliability can be found in Annex III.

3.5. Mode/Mechanism of action or adverse outcome pathways (MOA/AOP); including experimental (NAM) data and *in silico* models – e.g. prediction of MIEs, key events

There is yet no identified mode of action for developmental effects induced by valproic acid and analogue structures (i.e. a situation that will often be encountered in practice). The induction of the neural tube defect exencephaly by these compounds has been postulated to involve inhibition of HDAC enzymes as an important, possibly initial step (Ornoy, 2009; Lloyd, 2013), a mode of action that also appears relevant to humans (Wiltse *et al.*, 2005). However, there still is quite some uncertainty as to the exact AOP that is operational here

for these analogues: analogue accumulation within the foetus, oxidative stress, and folate antagonism are other possible suggested mechanisms (Lloyd, 2013).

With regard to HDAC inhibition as molecular initiating event, two postulated AOPs are submitted to AOPWiki, one for craniofacial development (#274), and one for neural tube defects (#275), both having HDAC inhibition as MIE. Linking *in vitro* assays to MIE and /or KE of these AOPs is not quite justified yet. Nonetheless, we have monitored histone deacetylase (HDAC) inhibition by the aliphatic carboxylic acids investigated in this study in our experimental models to verify the correlation between HDAC inhibition and exencephaly induction.

Because of this state of knowledge of neural tube defects underlying mode(s) of action (and developmental toxicity in general), and the state of knowledge of the applicability domains of the applied models, an attempt to link or select assays on this basis for this case study was considered not opportune.

The results are presented and discussed in chapters 4. and 5. , respectively.

3.6. Chemical/biological interaction

Section 3.8 below shows the OECD QSAR Toolbox profiler analysis results for the aliphatic carboxylic acids investigated in this study, that characterise the expected chemical-biological interactions for these chemical structures. Quite explicit characteristics are DNA and protein binding potential, but also potential for sensitisation, and irritation are reflecting expected chemical-biological interactions. From this analysis, it can be concluded that there are no great differences expected in the biological interactions of these category chemicals, which is another argument in support of their grouping.

3.7. Responses found in alternative assays (e.g., experimental (NAM) data, *in silico*)

The *in silico* responses are discussed in sections 3.6 and 3.8. For some of the chemicals investigated in this study quite some data in alternative assays is available in public literature. However, this data simply is too much to describe and discuss here. Moreover, a more realistic situation for most target chemicals is that this data is actually non-existent. Therefore, we simulate this case here, and only present and discuss the data we generate in this study: this is described in chapters 4. and 5. , respectively.

3.8. Information obtained from other endpoints

To further characterise, the biological similarity of the category members profilers of the OECD QSAR toolbox were used to gather information on DNA and protein binding, skin sensitisation, and skin and eye irritation.

Profiling for DNA and protein binding, and skin sensitisation

None of the grouped compounds has an alert for DNA and protein binding, and skin sensitisation. This is taken as similar behaviour of all compounds grouped in this study (Table 5). It is recognised though that the applied profilers flag the “presence of alerts” and that the absence of known alerts should not be interpreted as a lack of toxicity. As such, “negative predictions” should be taken cautiously.

Table 5. OECD Toolbox profiler predictions for DNA and protein binding, and on skin and eye irritation

CAS-RN Abbreviated name	88-09-5 EBA	97-61-0 MPA	4536-23-6 MHA	149-57-5 EHA	99-66-1 VPA	31080-39-4 PHA
Chromosomal aberrations						
OASIS	-	-	-	-	-	-
DNA binding						
OASIS	-	-	-	-	-	-
OECD	-	-	-	-	-	-
Sensitisation						
GHS	-	-	-	-	-	-
OASIS	-	-	-	-	-	-
h-CLAT	-	-	-	-	-	-
Protein binding						
OASIS	-	-	-	-	-	-
OECD	-	-	-	-	-	-
CYS potency (DPRA 13%)	-	-	-	-	-	-
GHS potency	-	-	-	-	-	-
LYS potency (DPRA 13%)	-	-	-	-	-	-
Skin irritation						
excl. BfR rules	u	u	u	u	u	Group C
incl. BfR rules	+	+	+	+	+	oa
Eye irritation						
excl. BfR rules	u	u	u	u	u	Group C
incl. BfR rules	oa	oa	oa	oa	oa	oa

- : no alerts, not reactive or unreactive; + : alert or reactive; oa: out of applicability domain

Profiling for skin and eye irritation

All compounds are flagged as irritating to skin, because they contain an aliphatic acid. This corresponds well to the pKa values of about 4.5 for all compounds. This alert is not indicating any dissimilarities between the target and source compounds.

Only PHA flags an alert for the exclusion rules for eye/skin irritation (i.e. this molecule should not be an irritant). It must be noted that such an alert is based on an estimated melting point and it is known that QSPR methods are characterised by prediction errors well in excess of the error of experimental measurements. Therefore, the uncertainty attached to the pertinence of such an alert is rather high (Dearden *et al.*, 2013).

From this data, it can be concluded that MHA and its analogues on both sides have no predicted dissimilarities in reactivity profile. All this data is also included in the data-matrix of Annex I (folder '*In chemico*').

3.9. Information on fate in the environment (hydrolysis, biodegradation)

Not applicable for the read-across case described here.

3.10. The route and duration of expected exposure.

Not applicable for the read-across case described here.

4. Data gap filling and Justification

In this chapter the *in vitro* and *in silico* methods used and results generated with these methods within the context of this read across approach for MHA will be shortly described.

Section 'Methodology' will name and describe the methods, as well as their suitability regarding the defined purpose of this study. Further details on these methods is described in referenced annexes at the back of this report. The section 'Results' will describe the results that have been generated with these described methods. Further details on these results is described in referenced annexes at the back of this report. The section 'Justification' will describe how the generated data are interpreted within the context of this study of assessing the developmental properties of MHA via reading across from structural analogues. This section will also evaluate the uncertainties of the data generated.

4.1. Methodology

In this section we will first of all describe the battery of experimental models applied in this study, followed by *in silico* models for prediction of biological activity, *in silico* models of toxicokinetic behaviour, and finally models for interpretation and classification of the results.

4.1.1. In vitro DART battery models

The experimental models applied are the ZET, the ZET reporter, the mEST and UKN1 models, and the CALUX reporter models. The suitability of the model and the applied scoring approaches will be outlined as well.

ZET

Suitability

The zebrafish embryo (*Danio rerio*) holds great promise as a model for developmental toxicity testing. The model is emerging as a candidate lower vertebrate animal model capable of filling the gap between high(er) throughput *in vitro* cellular assays and conventional preclinical animal testing. The embryonic/larval zebrafish model offers an intact whole animal model, providing the possibility to obtain information on dysmorphic effects comparable to the ones that may occur in rodent *in vivo* studies, with many of the advantages of *in vitro* systems, making it a suitable model organism for medium/high throughput screening (Ali *et al.*, 2011).

The model covers a broad developmental period (from fertilisation to hatching and early larval stage), including processes as gastrulation and neurulation and the development of major organ systems. It is relatively easy to monitor the morphology and evaluate effects on developmental since the embryos are transparent. In addition, for many years the zebrafish is a model to study (basic) embryology, so the development of the embryo is well defined (Kari *et al.*, 2007).

During these early developmental stages, the zebrafish embryos are not defined as protected under European legislation for animal welfare, as they are still yolk-dependent and thus non-free-feeding. In addition, fertilisation and development of the zebrafish occur

outside the body (oviparous), there is no need to sacrifice parental animals to obtain the embryos.

Method

Basically the method followed is that described in OECD guideline 236 and DB-ALM Fish Embryo Test 2013. Details are described in Annex II.

As deviation from this OECD 236 guideline the following is noted: as described (Braunbeck *et al.*, 2015; Braunbeck and Lammer, 2006; Embry *et al.*, 2010; Lammer *et al.*, 2009) embryos are raised (and exposed) until an age of 96 hpf. However, based on the wording of current EU animal welfare legislation (Directive 2010/63/EU of the European Parliament and of the counsel on the protection of animals used for scientific purposes), in cases of inconclusive observations, exposure may be extended to 120 hpf (Strähle *et al.*, 2012). In order to classify specific endpoints of the test substances, this extension has been performed in this study.

Calculation of EC/LC values and generating of graphs: EC and LC values were calculated with ToxRat Prof. Vers. 2.10. It also generated graphs for every observed effect, however, for a clear presentation the graphs of Annex II were created and edited with SigmaPlot 13.0. For the calculation of the EC/LC values and for the data of the graphs ToxRat Prof. Vers. 2.10, the total amount of affected embryos of every test concentration at every time point was used.

HDAC inhibition assay: the standard protocol for HDAC inhibition measurements applied in this project did not work out for zebrafish embryo's; one explanation may be that different HDAC isozymes are involved in this species.

Scoring

After the end of the test at 120 hpf, the number of embryos exhibiting specific effects were scored. Subsequently, from each dose effect graph EC_{10} values were calculated using ToxRat Professional version 2.10 and SigmaPlot 13.0 software. The EC_{10} was defined as the concentration at which there was a 10% increase of the incidence of a monitored effect, relative to control. Subsequently, the calculated EC_{10} values of all compounds were compared for ranking the compounds for their potency to induce the monitored effects.

ZET Reporter assay

Suitability

The suitability of the ZET has already been described above. This reporter assay focusses on one specific aspect of the developing zebrafish larvae: the deterioration of facial cartilage structures, called cranio-facial deformation.

Method

Basically, the method followed is that described in OECD guideline 236 and DB-ALM Fish Embryo Test 2013. Test solutions are refreshed every 24 hours. The measurement takes place at 120 hpf. The method relies on a transgenic zebrafish collagen reporter line, which expresses the fluorescent protein mCherry under the control of the collagen 2 a (col2a) promoter (Hammond and Schulte-Merker, 2009), and on the uniform positioning of larvae for imaging afforded by the Vertebrate Automated Screening Technology (VAST) bioimaging platform (Union Biometrica, USA). The basic raw data is fluorescent

ventral-view images of zebrafish larvae at 5 days post fertilisation. The images are analysed by measuring the angle formed by the ceratohyal bone (CHA), a characteristic v-shaped structure clearly visible in ventral views of 5 DPF larvae, as the increase in this angle provides a quantifiable proxy for the deterioration of the morphological development of facial cartilage structures. For more details, see Annex II.

HDAC inhibition assay: the standard protocol for HDAC inhibition measurements (applied to mEST, and UKN1 cells) did not work out for zebrafish embryo's; one explanation may be that different HDAC isozymes are involved in this species.

Scoring

After the end of the test at 120 hpf, the embryos are monitored for the CHA value. Thus, the basic analysis unit is degrees of angle. Data normalisation is based on the normalisation to mean measurements in sibling controls, defined as 0% response, while 100%, the complete deterioration of the ceratohyal angle, is defined as 180° where the two halves of the ceratohyal forms a straight line (see Annex II). The EC_{10} was defined as the concentration at which there was a 10% deviation of this CHA value relative to control. Subsequently, the calculated EC_{10} values of all compounds were compared for ranking the compounds for their potency to induce the monitored effects.

mEST

Suitability

The mouse embryonic stem cell test (mEST) is used as an *in vitro* model for the screening of embryotoxicity, based on a blastocyst-derived permanent embryonic mouse D3 cell line (ESC), derived from mouse 129 strains and was validated by the European Centre for the Validation of Alternative Methods (ECVAM). During this formal validation, the results obtained with the mEST concordances between the embryotoxic potential derived from the *in vitro* data and from the *in vivo* data were good, and were reproducible, both within and among different laboratories that performed the test. The ability of the test to discriminate a diverse group of chemicals, being strongly, weakly embryotoxic or non-embryotoxic, has been clearly demonstrated, and a predictivity of 100% was obtained with strongly embryotoxic chemicals (Genschow *et al.* 2002, 2004).

Method

The assay execution and data evaluation was performed according to the process described in the DB-ALM protocol provided in Annex II. The assay principle is based on the assessment of chemical-induced cytotoxicity and the inhibition of differentiation from mouse embryonic stem cells into cardiomyocytes during a ten days substance treatment.

The assay is defined through the three following endpoints: Inhibition of growth (cytotoxicity) of 3T3 fibroblast cells, which represent differentiated cells, and growth inhibition of undifferentiated ESC cells after 10 days of treatment. This is determined by the use of dehydrogenase enzymes present in the intact mitochondria of living cells to convert yellow soluble substrate 3-(4,5-dimethythiazol-2-yl)-2,5-diphenyl Tetrazolium bromide (MTT), into a dark blue insoluble formazan product, which gets sequestered within the cells and is detected quantitatively at 570 nm using a absorbance reader, after solubilizing the cell membrane. For this IC_{10} values are derived, i.e. inhibitory concentration of cell viability at 10% of the maximum dose response.

A third endpoint is the inhibition of differentiation of ESCs (embryonic stem cells) into cardiomyocytes after 10 days of treatment. The beating of these cells is evaluated by microscopy. Here the ID_{10} is derived, i.e. inhibitory concentration of differentiation at 10% maximum dose response.

HDAC inhibition assay: short description (see Annex II for details): ca 2.5 million D3 cells were lysed and the lysate containing the HDAC enzymes was then centrifuged for 10 mins at 10,000g. Supernatant was collected and diluted 1 in 6 with assay buffer. Reaction plate setup with the inhibitors (VPA analogues) and positive control (TSA) and incubated for 1hr at 37 degrees

Reaction stopped with a 1 μM TSA solution and then incubated for 15 minutes. Thereafter, the plate is read in the Tecan infinite M200 PRO reader at 360nm excitation and 460nm emission wavelength.

Scoring

From the differentiation assay and the cytotoxicity assays, half maximum dose response values were calculated (ID_{50}, and IC_{50} values resp.) and entered into a statistical evaluation developed from the modified prediction model used by Scholz *et al.* (1999), from which a prediction score was obtained to characterise chemicals as negative (D12_3 < 0.5), or positive (D12_3 > 0.6) with regard to embryotoxicity: see Annex II for details. Besides, relative ID_{10} values of chemicals were evaluated to assess their relative differentiation inhibition potential.

UKN1

Suitability

UKN1 model is a human neural embryonic stem cell test, corresponding to weeks 3-4 of human foetogenesis, to assess neurodevelopmental toxicity induction by chemical or drug exposure, and established with positive controls, classifiers for compound classes, biomarkers, extensive transcriptomics background data and epigenetic endpoints (Krug *et al.* 2013; Waldman et al 2014).

Method

UKN1 models the process of neural induction and patterning. In order to model this process human induced pluripotent stem cells (iPSC) are differentiated by dual SMAD inhibition for 6 days (Balmer, 2012; Krug, 2013). The differentiation protocol is based on previously published protocol (Chambers, 2009) and a detailed test method description and SOP is provided in Annex II.

The differentiated cells on day 6 correspond to neuroectodermal progenitors (NEP) as present during neural plate and neural tube formation. Therefore, disturbances of differentiation and patterning can be modelled by the test method. To this end, the differentiating cells are exposed to the compounds during the whole process of differentiation from day 0 to day 6. Neural tube defects are thought to be induced by wrong differentiation tracks during neurulation and in chicken; it has been shown that TSA-induced neural tube defects are due to induction of neural crest cell fade instead of NEP (Murko, 2010; Murko, 2013).

Therefore, the idea of the present test method is that we may detect neural tube defect inducing chemical by following gene expression changes. This includes e.g. spina bifida,

anencephaly or encephalocele. The scientific rational of the test method is based on the hypothesis that gene expression alterations induce a wrong differentiation of NEP cells or a changed differentiation track. For example if the iPSC cells do not differentiate into epithelial cells but in mesenchymal cells, the neural tube cannot be closed and severe congenital malformations may occur.

The UKN1 test method is mainly related to AOP #275 (implemented in the frame of the EUToxRisk project): "HDAC inhibition leads to neural tube defects" and measures two key events of this AOP. The first KE is "alterations of gene expression" and this is measured on day 6 by real time quantitative PCR (RT-qPCR) in order to follow the expression of 6 key player genes having a major role in the maintenance of pluripotency (OCT4, NANOG), neural induction (PAX6, OTX2) and a neural crest marker (TFAP2) which indicates wrong differentiation. A compound is counted as hit if PAX6 and OTX2 are downregulated and TFAP2 is upregulated. The second KE measured by UKN1 is "altered differentiation". To this end, we investigate if the NEP can be differentiated into the next developmental stage, which is represented *in vitro* by the so-called rosettes. The compound is removed on day 6 and the cells get further differentiated for another 8 days. On day 11, the cells are seeded in 96-well plates and are stained with two markers (ZO1, GM130) that mirror the specific morphology of the rosettes (Waldmann, 2014). The cells are imaged by an automated microscope and the number of rosettes per well is determined by an algorithm that segments the rosettes and counts the number of rosettes per well. If the number of rosettes is reduced the compound is counted as a hit. If both KE are affected by a compound we conclude on a DNT potential of the compound.

To determine the testing concentration for the measuring of KE1 and KE2 a viability based concentration response curve is measured and the EC_{10} is determined by a 4 parameter model fit. For the measuring of KE1 and KE2 (as described above) the ½ x EC_{10}, EC_{10} and 2x EC_{10} was used.

HDAC inhibition assay: short description (see Annex II for details): ca 2.5 million D3 cells were lysed and the lysate containing the HDAC enzymes was then centrifuged for 10 mins at 10,000g. Supernatant was collected and diluted 1 in 6 with assay buffer. Reaction plate setup with the inhibitors (VPA analogues) and positive control (TSA) and incubated for 1hr at 37 degrees.

Reaction stopped with a solution containing 1 µM TSA trypsin, and then incubated for 15 minutes. Thereafter, the plate is read in the Tecan infinite M200 PRO reader at 360nm excitation and 460nm emission wavelength.

Scoring

The prediction model to classify the compound as positive or negative is as follows:

1) PAX6 and OTX2 gene expression must be down-regulated (compared to control) and AP2 up-regulated.

2) The means of the fold change values are calculated (with down-regulated values are multiplied with -1) = gene expression score

3) The rosettes formation are transformed from % of control (which is set to 100%) to a scale from 1 (no inhibition of rosettes formation) to 20 (no rosettes are detectable) = RoFo score. The rational behind this scaling is that in this data set the maximal gene regulation was 17.5 fold. Therefore, as the RoFo score should

be weighted in the same range as gene expression, we have chosen 20 for maximal inhibition of rosettes formation.

4) The mean of the gene expression score and the RoFo score is then calculated = UKN1 score

The average UKN1 score of positive controls (known *in vivo* DNT effect) was 15.7 (SD 2.6) and 4.7 (SD 3.4) for the negatives. Based on SD, the following 3 classes were defined:

UKN1 score	classification
< 7	no hit (='0' in the data matrix)
7 – 13	no or weak hit (= 0.5 in the data matrix)
> 13	hit (= 1 in the data matrix)

CALUX reporters

Suitability

The CALUX panel is based on the U2-OS osteosarcoma cell line. This cell line does not endogenously express most nuclear receptors and/or metabolic enzymes. Therefore, this cell line is ideally suited to study the interaction of a compound with a specific receptor or pathway, without interference of metabolism or receptor cross-talk. This makes interpretation of the results relatively straightforward, and allows comparison of compounds on the level of MIEs. The CALUX panel consist of 26 cell-based reporter gene assays. Each cell line measures the activation or inhibition of one specific nuclear hormone receptor or cell signalling pathway (see below Figure 4).

In general CALUX assays focus on molecular initiating and early key events that are the primary target of toxicants (van der Burg *et al.*,2013; Becker *et al.*,2015), and involved in cellular events key in cell growth and organismal development, while deregulation can cause developmental disorders and diseases like cancer (NRC, 2000; Hanahan and Weinberg, 2011).

The panel is unique, in that it contains highly selective human cell based assays having a particularly strong coverage of assays for nuclear hormone receptors and pathways relevant for developmental- and reproductive toxicity (van der Burg *et al.*,2013; Sonneveld *et al.*,2005, 2006, 2011; Piersma *et al.*,2013; van der Burg *et al.*,2015a; Lewin *et al.*,2015), including a broader coverage within developmental and reproductive toxicity, but also covering additional relevant nuclear receptors (glucocorticoid-, retinoid-, peroxisome proliferator activated-, dioxin-, pregnane X-, liver X-, and progesterone receptors). In addition key pathways assays have been added including reporter gene assays for AP1 (fos/jun complexes, activated via MAPK pathways), NF-kappaB, TCF, nrf-2, p21 and p53 (van der Burg *et al.*,2013; Piersma *et al.*,2013; van der Linden *et al.*,2014; Gijsbers *et al.*,2013). Clearly most of these pathways are involved in quite generic responses including developmental and reproductive toxicity. The assay panel has been used to predict developmental toxicity successfully, either alone or in combination with other assays (Piersma *et al.,* 2013; Van der Burg *et al.*, 2015a,b). Also, a control cell line, called cytotox CALUX has been generated which constitutively expresses the same luciferase gene under control of the same expression plasmid which is used to generate reporter gene assays for steroid receptors (i.e. pSG5). This assay has been proved to be a very suitable control for high throughput screening (HTS).

All CALUX assays have now been automated in high throughput format. Data storage and analysis has been set up and more than 600 compounds have been screened in the CALUX

panel. Finally, the CALUX HTS panel endogenously expresses little metabolic activity but can be run with and without S9 metabolic fractions, allow assessment of involvement of metabolism in chemical effects (van Vugt-Lussenburg *et al.*, 2018).

The assay results can be used to identify compounds activating similar MIE's, and is therefore well suited for read across studies. Furthermore, the results provide clues for the MoA of the different compounds. Therefore, a direct link with an AOP (if available) can be made.

Role of HDAC inhibition on activation profile CALUX reporters

HDAC is a positive regulator of gene expression in multiple settings. HDACs are recruited by a variety of transcription factor corepressor complexes and are believed to repress transcription by reducing the level of acetylation of core histones, thereby altering chromatin structure. HDAC inhibition, therefore, will result in an increase in activity of several reporters and cell signalling pathways. The CALUX activity profile of chemicals could also be explained by their activity as HDAC inhibitors, rather than direct interaction with the pathways (Piersma *et al.* 2013).

Method

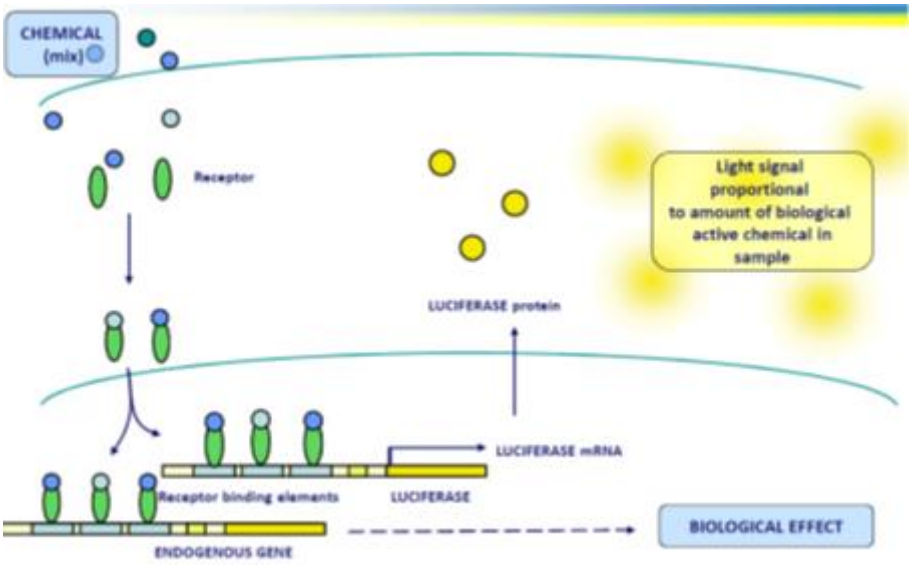
The CALUX panel of 26 cell based reporter gene assays are depicted in the table below.

Table 6. The CALUX panel used in this Case Study.

Cell line	Endpoint	Cell line	Endpoint
Cytotox CALUX	Cytotoxicity	PPARg CALUX	Peroxisome proliferator receptor agonists
ERa CALUX (ago/anta)	Estrogen receptor (ant)agonists	AhR CALUX	Aryl hydrocarbon receptor agonists
AR CALUX (ago/anta)	Androgen receptor (ant)agonists	Hif1a CALUX	Chemical hypoxia response
PR CALUX (ago/anta)	Progesterone receptor (ant)agonists	TCF CALUX	Wnt/TCF pathway activation
GR CALUX (ago/anta)	Glucocorticoid receptor (ant)agonists	AP1 CALUX	AP1 pathway activation/cell cycle control
TRb CALUX (ago/anta)	Thyroid receptor (ant)agonists	ESRE CALUX	Endoplasmic reticulum stress
RAR CALUX	Retinoid acid receptor agonists	NFkB CALUX	Activation of NF-kB pathway (immune response)
LXR CALUX	Liver X receptor agonists	Nrf2 CALUX	Oxidative stress
PXR CALUX	Pregnane X receptor agonists	p21 CALUX	Transcription of p21 inhibitor of cell cycle progression
PPARa CALUX	Peroxisome proliferator receptor agonists	p53 GENTOX CALUX	P53-dependent pathway activation/genotoxicity
PPARd CALUX	Peroxisome proliferator receptor agonists		

CALUX assay principle used for this read across study is shown below in Figure 5. For the current study, DB- ALM Protocol n° 197: Automated CALUX reporter gene assay procedure was used; this method is described in the combined OECD 211/ DB-ALM template (Annex I). This involves exposure to the study compounds in a concentration range of 1E-3M (1 mM, as maximum tested concentration) to 1E-9M (1 nM, as lowest tested concentration) of 26 different CALUX assays, carried out by a liquid handling robot.

Figure 5. CALUX assay principle.



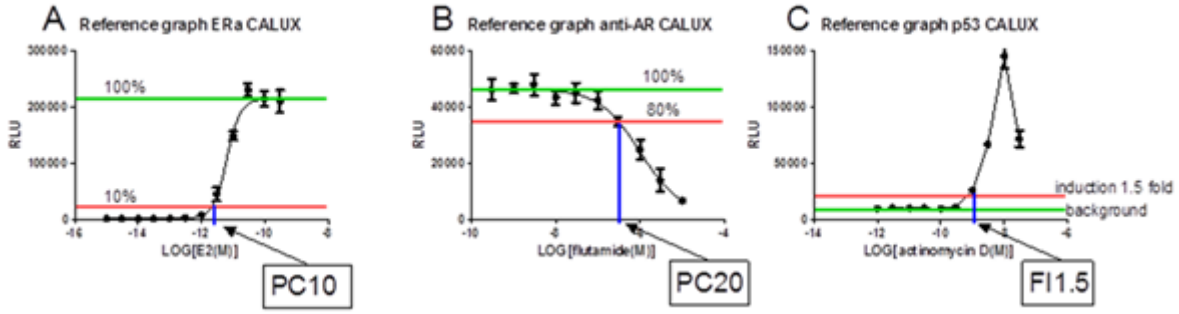
In 2005 we already developed a panel of mechanism-based CALUX assays to assess hormonal activity of compounds (Sonneveld *et al.*, 2005), a panel which has shown to be highly predictive for such activities in experimental animals (Sonneveld *et al.*, 2006, 2011). Starting in 2005, several of these assays have been successfully engaged in extensive validation exercises as alternatives to animal experiments via EURL-ECVAM, OECD, and others (OECD, 2009, 2013a, 2013b, van der Burg *et al.*, 2010a, 2010b). All assays are validated in house and at client laboratories. Further formal validations have been initiated in the area of thyroid disruption, metabolic activation of endocrine assays, while SOPs, including one on the assay panel automation, have been submitted to the EURL-ECVAM database of alternative methods DB-ALM.

Scoring

Reported values are lowest effect concentrations (LEC) in Log(M). Depending on the assay, LECs are defined as (Figure 6, next page):

- the concentration where the test compound causes an **activation/agonist** effect equal to 10% of the maximum effect elicited by the test's reference compound (PC10, Figure 6A);
- the concentration where the test compound causes an **antagonist** effect equal to 20% of the maximum effect elicited by the test's reference compound (PC20, Figure 6B);
- the concentration where the test compound elicits **pathway activation** 1.5-fold above background (FI 1.5, Figure 6C)

Figure 6. Definition of lowest effect concentrations in mechanistically different CALUX assays.



4.1.2. In silico models

QSAR for prediction in vivo induction of exencephaly in rodents

The prediction model was built based on the transformation of the known qualitatively observed *in vivo* structure-activity features of VPA analogues associated with the induction of exencephaly into numerical factors. These structural features are combined with the HDAC inhibitory potency and relevant physical-chemical parameters, indicated to affect the pharmacokinetics of the respective analogue, to build a prediction model that gives the probability of a certain analogue to induce exencephaly. This QSAR approach is one of a type where a biological layer of data is added to chemical data to strengthen and cover-up for unconsidered parameters that affect the individual analogue activity *in vivo*. From the model, another output was the teratogenicity score that allows extrapolating and obtaining an analogue's teratogenicity potential on the Nau laboratory dose-exencephaly rate scale (from 0 to +++++). The prediction model is built using linear discriminant analysis leave-one out cross validation on a set of 28 valproic acid analogues with twenty two as the training set. The other analogues are then used to test the developed model and the predictions are cross-checked against recorded *in vivo* observations in the Nau-protocol based studies. A more detailed description of the method is provided in Annex III.

Toxicokinetic models

Bio-kinetic modelling

mEST, UKN1, and CALUX reporter assays

Effective concentrations determined in *in vitro* toxicological assays are routinely based on a range of nominal treatment concentrations. However, the use of the nominal treatment concentration as the driving concentration for observed toxicity *in vitro* does not account for factors that reduce the free concentration within the assay and determine the true effective concentration. For cell based assay systems, this will dictate the concentration available for distribution into the cell, and subcellular compartments, and so determine the driving concentration at the target site mediating toxicity. In order to more accurately translate concentration driven toxicity from *in vitro* to *in vivo*, it is necessary to correct the nominal effect concentration, accounting for the distribution of the compound within the assay system. Factors to be considered in modelling this *in vitro* distribution include binding to the plastics used in assays, exchange at the interface between culture media and

the air in the culture vessel, and binding to components that may be included in the culture media (e.g. lipids and proteins originating from foetal bovine serum, FBS); the modelling of these processes is termed biokinetics (Blaauboer *et al.*, 2010). Models developed for CALUX (termed 'Steady-state Biokinetic Model 2D Monolayer Cell Culture' model), and for mEST, and UKN1 (termed 'Alternative Biokinetic Predictions' model) are described in more detail in Annex III, and the relevant *in vitro* concentrations named 'cell-total', and 'medium unbound', respectively.

Zebrafish Embryo Model

To better explain, predict, and extrapolate developmental toxicity observed in zebrafish embryo, a PBPK model for zebrafish embryo was developed which integrates organ volume changes during development and hepatic metabolism. Our zebrafish embryo model is adapted from Péry *et al.* (2014) for adult zebrafish model and on the Simcyp® virtual *in vitro* intracellular distribution (VIVD) model (Fisher *et al.*, 2019). The model assumes quasi steady-state distribution between various zebrafish cell types (tissues). The partition coefficients between the various organs and organelles considered by the model (yolk, liver, gut, muscle, skeleton, eye, brain, heart, skin, other tissues, lysosomes, and mitochondria) and the culture medium, are conditioned by the chemical of interest physiochemical properties (pH acidic, basic, neutral or amphoteric character of the molecule, etc.). A more detailed description of the method is provided in Annex III. The relevant *in vitro* concentrations were named 'embryo-total' concentrations.

PBPK modelling

PBPK models were used to convert *in vitro* dose descriptors derived from the above mentioned models into Oral Equivalent Doses (OED) by reverse dosimetry, one model to calculate moose OEDs, and one for human OEDs. In this modelling both minimal and maximal macromolecular binding in plasma were explored as two extreme *in vivo* conditions critically determining this OED. A description of these models and used reasoning can be found in section 3.4, and further details are provided in Annex III.

4.1.3. Classification and data analysis models

Dempster-Shafer Theory approach

The Dempster-Shafer theory (Shafer G., 1976; Dempster AP, 1967; DST) is a promising regulatory tool when making decisions in the face of multiple, and sometimes even conflicting, lines of evidence (Park *et al.*, 2013), that was applied to the data in this case study. DST is an extension of generalised Bayesian statistical inference in which evidence can be associated with multiple sources.

DST represents a rigorous decision-theory approach that provides a framework to generate predictions, estimate the uncertainty associated with each prediction, and combine multiple sources of evidence resulting in a weight-of-evidence (WoE) prediction by quantitatively accounting for the reliability of each of the individual sources (details in Annex III). In general, this decision theory will support the decision making process done by the toxicologist.

In this submission, DST was used to combine the evidence from different *in vitro* assays for the source compounds in order to provide a WoE estimate for the *in vivo* neurodevelopmental toxicity outcome.

First the subset of assays were identified, which gives the best validation results from a leave-one-out (LOO) cross validation. This LOO validation enables to detect the reliability, positive prediction accuracy (PPV) and negative prediction accuracy (NPV) of the assays. The LOO procedure is a standard approach for selecting a set of sources (in this case assays) based only on the training set (in this case the source compounds).

Bayesian Automatic Classification approach

In addition to expert judgment on the meaning of the individual test results, an automatic classification method was used to decide, considering all the evidence, in which category the target compound is likely to fall. The predicted *in vivo* NTD was discretised in 5 levels (missing data, negative, mildly positive, positive, strongly positive. Only MPA and MHA had missing *in vivo* data (in fact MHA has *in vivo* data, but they were blinded to the analysis to check how MHA would be predicted). A Bayesian consensus approach was developed. The method can be summarised as follows:

- a prediction model was fitted to individual test data on training chemicals (source chemicals for which both *in vitro*-derived effect measures and *in vivo* data were available); A Bayesian framework and Bayesian numerical calibration algorithms (MCMC sampling) were used, under normality assumption for the *in vivo* data likelihood, given the *in vitro* predictors (those key-event measurements were continuous variables);
- the poorly predictive tests were removed from consideration on the basis of posterior predictive checking of their classification accuracy (on the training set of chemicals only, we had too few chemicals to do a formal validation);
- use the remaining (sufficiently predictive) tests to build a set of predictions for the test chemical (MHA or MPA).

Biological Fingerprint Classification approach

For this statistical analysis, the *in vitro* results (continuous variables) were converted into a binary coding (active/non-active) to obtain fingerprints per chemical. Subsequently, from these fingerprints chemical distances (Tanimoto scoring) were obtained, expressing *in vitro* response profile similarity of couples of chemicals. From this dendrograms were drafted, and clustering of chemicals was examined.

4.2. Results

In this section, we will describe the results that have been generated with the above described *in vitro* and *in silico* methods. All *in vitro* models delivered nominal EC_{10} values for the respective read outs monitored. As no clear mode of action or even AOP has been identified for the effects induced by the source chemicals, all available model read outs have been monitored for EC_{10} or comparable dose descriptor values. These nominal EC_{10} values were converted into biologically more relevant EC_{10} values by the TK models described in section 4.1.2 (and further detailed in Annex III): for Zebrafish-embryo into 'total embryo' EC_{10} values, into 'unbound medium' ID_{10}, IC_{10}, and EC_{10} values, for the cellular mEST, and UKN1, respectively, and into 'cell-total' LEC values for CALUX reporter assays. These biologically more relevant EC_{10} values represent relative intrinsic potencies of the chemicals to induce effects in these various *in vitro* systems. Some of these systems do have 'objective' classification scores, however, we will assess the relative

potency values of these structurally related chemicals as they are considered critical in assessing the *in vivo* hazard potential.

For 5 of these compounds, i.e. for EBA, MPA, MHA, EHA, and VPA, PBPK models were developed to calculate – via reverse dosimetry – oral equivalent doses (OEDs) that correspond to the derived biologically relevant dose descriptors for ZET ('total embryo' EC_{10} value), for mEST ('unbound medium' ID_{10}, and IC_{10} values), for UKN1 ('unbound medium' EC_{10} value), and for CALUX reporters (cell-total LEC). This reverse dosimetry is explained in the toxicokinetic modelling section 3.4.

It is again emphasised that of the 9 compounds tested in the *in vitro* models in this study 5 have asymmetric carbon atoms, and thus may exist as pure R- or S- enantiomers: MPA, MHA, EHA, PHA, and 4-ene-VPA. In this study, we have tested the racemates, herewith aligning to the *in vivo* data that we refer to, as these were generated with the racemates as well.

4.2.1. In vitro DART battery models

ZET

All data are described in detail in Annex II, while in the data-matrix (Annex I) and the below Table 7 describe the EC_{10} values. For 4-ene-VPA, unfortunately, experimental results are lacking.

Table 7. Nominal and total embryo EC_{10} values (in µM) in ZET.

Abbreviated name		PA	DMPA	EBA	MPA	MHA	EHA	VPA	PHA
Most sensitive effect	*nominal*	568	426	363	373	385	65	38	10
	total embryo	142	107	61	93	96	13	8	1
Small eyes	*nominal*	>1400[1,3]	540	(1000)[2,3]	>900[3]	>1000[3]	(400)[3]	(400)[3]	153
	total embryo	>350	135	(167)	>225	>250	(80)	(80)	19
Jitter/tremor	*nominal*	>1400[3]	977[3]	>1000[3]	>900[3]	442	421[3]	82	16
	total embryo	>350	244	>167	>225	111	84	16	2
Cranio-facial deform.	*nominal*	626[3]	481	>1000[3]	380	>1000[3]	102	52	21
	total embryo	157	120	>167	95	>225	20	10	3
Scoliosis/lordosis	*nominal*	747[3]	663[3]	892[3]	393	383	651[3]	59	41
	total embryo	187	166	144	98	96	130	12	5

1) '>'means: not detectable up to this highest tested dose; 2) '(..)' means: at highest tested dose <10% effect observed; 3) recently[1], it was noticed that for this dose the medium pH value will have dropped to below pH 6.5 (the OECD Guideline 236 indicated border), and the depicted total embryo EC_{10} value will be an underestimation of the true embryo value.

In vitro total embryo EC_{10} values

The most sensitive endpoint for most of the analogues was pericardial oedema; PHA, VPA and EHA showed most potent in inducing yolk oedema or yolk darkening. This effect shows the embryo is starting to experience toxic effects; these effects mostly are still considered reversible, though. These effects occurred in embryos treated with PHA,VPA,

[1] This was discovered late June 2020; legend notes on this to Tables 7 to 10 were added August 2020 after the review of the case study; the authors conclude that it will impact on some of the derived quantitative results, i.e. increasing some of the higher EC_{10} and OED values, but it will not impact the concept and derived conclusions of this read-across case study.

and EHA at clearly lower concentrations than with the other analogues, PHA being most potent, followed by VPA, and EHA, with total embryo EC_{10} values of 1, 8, and 13 µM, respectively. The other category analogues EBA and MPA had higher total embryo EC_{10} values starting from 61 µM, while DMPA and PA had values of 107 µM and higher.

When looking at the neurotoxic effects jitter/tremor and small eyes, PHA, and VPA appeared most potent, with total embryo EC_{10} values of 2, and 16 µM, respectively. The category analogue MPA didn't show these effects, while EHA, MHA, and EBA scored in between with total embryo EC_{10} values of 84, 111, and >167 µM[2] , respectively. PA didn't show these effects, while DMPA had a total embryo EC_{10} values of 135 µM.

Clear developmental effects like cranio-facial deformation or scoliosis & lordosis showed a trend similar as seen with the most sensitive effects: i.e. these effects were induced by PHA, VPA, and EHA at clearly lower concentrations as with all other analogues, total embryo EC_{10} values being 3, 10, and 20 µM, respectively, while these effects occurred at total embryo EC_{10} values of 95, 96, and 144 µM, for MPA, MHA, and EBA, respectively, and 120, and 157 µM, for DMPA, and PA, respectively.

Corresponding mouse and human OED values

Mouse and human OED values for the respective effects monitored are depicted below in Table 8 for the five compounds for which PBPK models were available. The OED values correspond to the indicated *in vitro* specified effect concentrations, i.e. it is the relative external dose needed to achieve such an internal equipotent target concentration.

Table 8. ZET total embryo EC_{10} values (in µM), and corresponding OED values (µmol/kg) in mouse and human.

Abbreviated name		EBA	MPA	MHA	EHA	VPA
Most sensitive effect	*ZET total embryo*	61	93	96	13	8
	Mouse OED range	490-1060	730-1470	210-360	60-100	3-7
	Human OED range	440-460	530-730	420-460	30-40	20-30
Small eyes	*ZET total embryo*	(167) [1,3]	>225 [2,3]	>250[3]	(80)[3]	(80)[3]
	Mouse OED range	(1340-2890)	>1770-3570	>550-880	330-610	50-80
	Human OED range	(1210-1260)	>1700-1760	>1100-1150	210-260	190-250
Jitter/tremor	*ZET total embryo*	>167[3]	>225[3]	111	84[3]	16
	Mouse OED range	>1340-2890	>1770-3570	250-400	340-640	9-15
	Human OED range	>1210-1260	>1700-1760	490-510	220-270	39-50
Cranio-facial deform.	*ZET total embryo*	>167[3]	95	>250[3]	20	10
	Mouse OED range	>1340-2890	750-1500	>550-880	80-150	6-9
	Human OED range	>1210-1260	710-740	>1100-1150	50-60	24-31
Scoliosis/lordosis	*ZET total embryo*	144[3]	98	96	130[3]	12
	Mouse OED range	1160-2490	770-1550	210-360	530-990	7-14
	Human OED range	1040-1090	740-770	420-460	330-420	28-35

1) '(..)' means: at highest tested dose <10% effect observed; 2) '>'means: not detectable up to this highest tested dose; 3) recently[3], it was noticed that for this dose the medium pH value will have dropped to below pH 6.5 (the OECD Guideline 236 indicated border), the depicted total embryo EC_{10} value and OED ranges will be underestimations of the true embryo and OED values.

[2] '>': means no observed effects at highest tested dose.

[3] See footnote 1, page 45.

When looking at the mouse OED ranges, the following observations are made. For the most sensitive endpoint pericardial and yolk oedema, both MHA and EHA are somewhat in between EBA, and MPA on the one hand (highest OED values), and VPA on the other hand (lowest OED value): EHA is closer to VPA, while MHA is closer to EBA, and MPA.

For neurodevelopmental effects, and for developmental effects, again MHA and EHA had somewhat in between potencies of inducing these effects, as compared to VPA on the one hand being most potent, and EBA, and MPA on the other hand being least potent, if at all capable of inducing these effects. Potencies of EHA and MHA were more close here (as compared to most sensitive effects), EHA being somewhat more potent in inducing small eyes, and cranio-facial deformation, while MHA appeared slightly more potent in inducing jitter/tremor, and scoliosis/lordosis. Please, note that the '>' sign on front of the OED values are to indicate that these were highest tested doses at which these effects have not been observed; these values were generated for comparison purposes only.

For human OED ranges, the picture is partly different. MHA clusters for the most the sensitive endpoint pericardial and yolk oedema with EBA, and MPA (i.e. has comparable OED values), while EHA and VPA cluster here separately (having clearly lower OED values).

For the neurodevelopmental endpoints, MHA again clusters with EBA, and MPA for small eyes, while EHA clearly clusters with VPA, but for jitter/tremor, MHA has a somewhat lower OED than EBA, and MPA, but still about an order of magnitude higher than VPA, while EHA has an OED value half that of MHA. For the developmental effects cranio-facial deformation and scoliosis/lordosis a somewhat similar picture as for neurodevelopmental effects emerges. Again, please, note that the '>' sign on front of the OED values are to indicate that these were highest tested doses at which these effects have not been observed; these values were generated for comparison purposes only.

ZET Reporter

All data are described in detail in Annex II, while EC_{10} values and OEDs are depicted in Annex I, and in the below tables.

In vitro total embryo EC_{10} values

From this data it appears that this assay is much more sensitive for monitoring cranio-facial deformation than the above described model from Heidelberg; this may be due to a strain difference, and difference in methodological sensitivity. VPA and PHA already induced cranio-facial deformation at relatively low total embryo EC_{10} concentrations of 0.1, and 0.3 µM, respectively. EHA is somewhat less potent with a total embryo EC_{10} of 2 µM, while MPA and EBA are only effective at much higher concentrations of 49 and 143 µM. For MHA no cranio-facial deformation was observed up to the highest tested total embryo dose of >20 µM; higher doses resulted in clear decrease in survival. The 'positive control' 4-ene-VPA had a total embryo EC_{10} of 1 µM for this effect, PA showed completely inactive (total embryo EC_{10} >500 µM), while DMPA was equally potent as MPA, with a total embryo EC_{10} value of 49 µM. All this is shown in Table 9 below.

Table 9. Nominal and total embryo EC_{10} values (in µM) in ZET CHA Reporter assay.

Abbreviated name		PA	DMPA	EBA	MPA	MHA	EHA	VPA	PHA	4-e-VPA
CHA effect	*nominal*	>2000 [1,2]	174	856[2]	194	>80	12	2	1	7
	total embryo	>500	44	143	49	>20	2	0.5	0.1	1

1) '>'means: not detectable up to this highest tested dose; 2) recently[4], it was noticed that for this dose the medium pH value will have dropped to below pH 6.5 (the OECD Guideline 236 indicated border), and the depicted total embryo EC_{10} value will be an underestimation of the true embryo value.

Corresponding mouse and human OED values

Mouse and human OED values for this CHA effect are depicted below in Table 10 for the five compounds for which PBPK models were available. The OED values correspond to the indicated *in vitro* specified effect concentrations, i.e. it is the external dose needed to achieve such an internal equipotent target concentration.

Please, note that no CHA effect was observed for MHA up to the highest tested dose (indicated by the '>'sign), which is a relatively low dose as at higher dose levels survival was substantially reduced. The associated OED values were generated just for comparison purposes.

Table 10. ZET reporter total embryo EC_{10} values (in µM), and corresponding OED values (µmol/kg) in mouse and human.

Abbreviated name		EBA	MPA	MHA	EHA	VPA
CHA effect	*ZET total embryo*	143 [2]	49	>20 [1]	2	0.5
	Mouse OED range	1154-2472	388-775	>44-71	7-15	0.14-0.28
	Human OED range	1034-1077	370-388	>88-92	5-7	0.73-0.94

1) '>'means: not detectable up to this highest tested dose; 2) recently[4], it was noticed that for this dose the medium pH value will have dropped to below pH 6.5 (the OECD Guideline 236 indicated border), and the depicted total embryo EC_{10} value will be an underestimation of the true embryo value.

From this table one can see that the OED values for MHA and EHA are somewhat in between EBA, and MPA on the one hand (highest OED values), and VPA on the other hand (lowest OED value): EHA clearly is closer to VPA, while MHA is closer to EBA, and MPA; how close this is cannot be determined, as it was not observed at the indicated (highest tested) dose in this model.

mEST

Full dose response data description and standard analysis results are provided in Annex II, while relevant ID_{10} values for D3 cells, and IC_{10} values for cytotoxicity of D3 and 3T3 cells, as well as HDAC inhibition and OED data are depicted in the data-matrix (Annex I).

In vitro medium unbound ID_{10}, IC_{10} and EC_{10} values

For EBA MPA and the 'negative control' DMPA, no effects were observed on differentiation up to nominal concentrations of 3000 µM, and thus, medium unbound ID_{10} concentrations were not derived for them (see below Table 11). Medium unbound ID_{10} values for VPA, and PHA were 189, and 116 µM, respectively. EHA and MHA followed

[4] See footnote 1, page 45.

with medium unbound ID_{10} values of 406, and 875 µM, respectively. The 'positive control' 4-ene-VPA had a medium unbound ID_{10} value of 212 µM, while the other 'negative control' PA showed a medium unbound ID_{10} value of <78 µM.

Table 11. Nominal and medium unbound ID_{10}, IC_{10} and EC_{10} values (in µM) in mEST.

Abbreviated name		PA	DMPA	EBA	MPA	MHA	EHA	VPA	PHA	4-e-VPA
ID_{10} D3	*nominal*	72	>3000 [1]	>3000	>3000	913	547	275	278	328
	medium unbound	71	>2490	>2910	>2880	875	406	189	116	212
IC_{10} D3	*nominal*	<78	>3000	2064	2713	>10000	<313	415	294	343
	medium unbound	*70*	>2490	2000	2601	>9580	*300*	285	123	222
IC_{10} 3T3	*nominal*	79	>3000	>3000	1838	6834	9568	>3000	1084	354
	medium unbound	78	>2490	>2910	1762	6552	7101	>2060	452	229
HDAC EC_{10}, day 4	*nominal*	950	>10000	4640	7380	>9580	340	50	40	390
	medium unbound	935	>8290	4496	7076	>10000	252	34	17	253
HDAC EC_{10}, day 10	*nominal*	30	3680	2130	1500	2400	160	40	30	50
	medium unbound	30	3051	2064	1438	2301	119	27	13	32

1) '>'means: not detectable up to this highest tested dose.

The corresponding medium unbound IC_{10} values in D3 were as follows: 2000, and 2601 µM for EBA, and MPA, respectively, while not detectable for the 'negative control' DMPA, i.e. comparable or somewhat lower as compared to their ID_{10} values. Those analogues having the lowest medium unbound ID_{10} values, i.e. VPA, and PHA also had the lowest medium unbound IC_{10} values: 285, and 123 µM, respectively. Of the remaining two analogues, EHA and MHA, the medium unbound IC_{10} values were 232 µM, and undetectably high, respectively. The 'positive control' 4-ene-VPA had a medium unbound IC_{10} value of 222 µM, while the 'negative control' PA showed a medium unbound ID_{10} value of <78 µM.

For 3T3 cells, medium unbound IC_{10} values were as follows: for EBA, MPA and MHA, and the controls PA, DMPA, and 4-ene-VPA they were more or less comparable to the medium unbound IC_{10} values for D3 cells. For the other analogues, EHA, VPA, and PHA, these were considerably higher as compared to their corresponding values in D3 cells.

HDAC inhibition

Concerning HDAC inhibition in D3 cells the following results were observed. All compounds tested appeared to be more active when monitored at the 10th day of incubation. Still, EBA, MPA and MHA were virtually inactive in this respect: their medium unbound IC_{10} values were 2064, 1438, and 2301 µM, respectively. Medium unbound IC_{10} values of the other analogues EHA, VPA, and PHA were 119, 27, and 13 µM, respectively. The 'positive control' 4-ene-VPA had a medium unbound IC_{10} value of 32 µM, while the 'negative controls' DMPA and PA showed medium unbound IC_{10} values of 3051 and 30 µM, respectively. Please, note that HDAC monitoring used a different experimental protocol, and thus absolute dose descriptor values achieved here should not be compared with absolute descriptor values achieved for functional effects.

Corresponding mouse and human OED values

Mouse and human OED values for D3 differentiation, and HDAC inhibition are depicted below in Table 12 for the five compounds for which PBPK models were available. The

OED values correspond to the indicated *in vitro* specified effect concentrations, i.e. it is the external dose needed to achieve such an internal equipotent target concentration.

Table 12. mEST medium unbound ID_{10} and HDAC EC_{10} values (in µM), and corresponding OED values (mmol/kg) in mouse and human.

Abbreviated name		EBA	MPA	MHA	EHA	VPA
D3 ID_{10}	*medium unbound*	>3000 [1]	>3000	875	406	189
	Mouse OED range	>26.0-31.8	>25.2-30.1	2.7-2.8	3.8-4.1	0.21-0.27
	Human OED range	>4.3-4.6	>4.5-4.7	0.84-0.95	0.23-0.36	0.11-0.17
HDAC EC_{10}, day 10	*medium unbound*	2064	1438	2301	119	27
	Mouse OED range	18.52-22.60	12.58-15.01	7.12-7.32	1.12-1.19	0.028-0.042
	Human OED range	3.08-3.27	2.25-2.37	2.2-2.51	0.07-0.1	0.014-0.024

1) '>'means: not detectable up to this highest tested dose.

Both MHA and EHA are somewhat in between EBA, and MPA on the one hand (highest OED values), and VPA on the other hand (lowest OED value). The OED value for differentiation inhibition for MHA and EHA are quite close for mouse, while it seems to be somewhat lower for EHA for humans, close to that of VPA. For HDAC inhibition, OEDs for mouse and human are much lower for EHA as they are for MHA.

UKN1

Concentration response curves for viability are shown in Annex II, as well as an overview of graphs of expression of PAX6, OTX2 and AP2 genes, and the endpoint rosette formation at the specified viability concentrations. OED values are included in Annex I as well.

In vitro medium unbound EC_{10} values

The below Table 13 shows that cell viability reduction for EBA, MPA, and MHA is not observed up to the highest tested nominal dose of 5000 µM, which corresponds to medium unbound EC_{10} values of >4487, >4319, and >4321 µM, respectively. Unbound medium EC_{10} values for EHA, VPA, and PHA, are 184, 215, and 43 µM, respectively, i.e. much lower. The 'positive control' 4-ene-VPA had a medium unbound IC_{10} value of 121 µM, while the 'negative controls' DMPA and PA showed medium unbound IC_{10} values of 1164 and 1710 µM, respectively.

Table 13. Nominal and medium unbound EC_{10} values (in µM) in UKN1 model.

Abbreviated name		PA	DMPA	EBA	MPA	MHA	EHA	VPA	PHA	4-e-VPA
EC_{10} viability,	*nominal*	1807	2042	>5000	>5000	>5000	417	575	263	363
	medium unbound	1710	1164	>4476	>4319	>4321	184	215	43	121
HDAC EC_{10},	*nominal*	90	1240	1040	1190	2370	160	90	30	110
	medium unbound	85	707	931	1028	2048	70	34	5	37

At these EC_{10} values gene expression and rosette formation were as follows. Those analogues having relatively low EC_{10} values, i.e. EHA, VPA, PHA, and 4-ene-VPA, all had clear reductions in rosette formation, and the anticipated gene expression changes, with the exception of PHA: this analogue did not show the gene expression changes accompanying inhibition of rosette formation. Prediction scores were positive for EHA, VPA, and 4-ene-VPA (UKN1 scores all >16.3), while PHA was scored as weakly positive

(UKN1 score of 11.9). The compounds with the relatively high EC_{10} values of > 5000 µM, EBA, MHA, and MPA, all have somewhat different profiles for gene expression, and inhibition of rosette formation: while EBA has no effects in these two parameters, MHA shows some gene expression changes and inhibition of rosette formation. Finally, MPA shows some of the anticipated gene expression changes, but without inhibition of rosette formation. The associated prediction scores classified EBA and MPA as being negative, while MHA was classified as unclear/weakly positive (UKN1 score of 7.9). Of the medium potent analogues with regard to viability, DMPA, and PA (EC_{10} values resp. 2042, and 1807 µM, and UKN1 scores of 9.6, and 17.6 resp.), PA showed most clearly showed gene expression changes, and rosette formation inhibition. Based on these profiles PA was classified as positive, DMPA as unclear/weakly positive.

HDAC inhibition

Concerning HDAC inhibition (monitored at day 6) the following results were observed. EBA, MPA, and MHA had relatively high EC_{10} medium unbound IC_{10} values of 1040, 1190, and 2370, respectively. Medium unbound IC_{10} values of the other analogues EHA, VPA, and PHA clearly were lower. i.e. 160, 90, and 30 µM, respectively. The 'positive control' 4-ene-VPA had a medium unbound IC_{10} value of 110 µM, while the 'negative controls' DMPA and PA showed medium unbound IC_{10} values of 1240, and 90 µM, respectively. Please, note that HDAC monitoring used a different experimental protocol, and thus absolute dose descriptor values achieved here should not be compared with absolute descriptor values achieved for functional effects.

Corresponding mouse and human OED values

Mouse and human OED values for UKN1 cell viability, and HDAC inhibition are depicted below in Table 14 for the five compounds for which PBPK models were available. The OED values correspond to the indicated *in vitro* specified effect concentrations, i.e. it is the external dose needed to achieve such an internal equipotent target concentration. Please, note that the '>' sign on front of the OED values are to indicate that these were highest tested doses at which these effects have not been observed; these values were generated for comparison purposes only.

Table 14. UKN1 medium unbound cell viability EC_{10} and HDAC EC_{10} values (in µM), and corresponding OED values (mmol/kg) in mouse and human.

Abbreviated name		EBA	MPA	MHA	EHA	VPA
Cell viability EC_{10}	*medium unbound*	>4476	>4319	>4321	184	215
	Mouse OED range	>40-48	>38-45	>13.4-13.7	1.7-1.85	0.24-0.31
	Human OED range	>6.7-7.1	>6.8-7.1	>4.13-4.71	0.10-0.16	0.12-0.19
HDAC EC_{10}	*medium unbound*	931	1028	2048	70	34
	Mouse OED range	8.4-10.2	9-10.7	6.34-6.51	0.66-0.70	0.035-0.049
	Human OED range	1.39-1.47	1.61-1.70	1.96-2.23	0.04-0.06	0.021-0.028

EBA and MPA show highest OED values both for mouse, as well as for human, both for viability reduction, as well as HDAC inhibition. MHA has comparable OED values as EBA, and MPA, apart from mouse viability reduction, where its OED is an order of magnitude lower. VPA clearly has the lowest OED values of all, equalled by EHA, apart from mouse HDAC inhibition, where its OED is an order of magnitude higher. All OED values are clearly below those of MHA, mostly an order of magnitude.

CALUX Reporters

In vitro cell-total LEC values

Assays important for the endpoint under investigation, i.e. induction of neural tube defects, are expected to show a positive response to the *in vivo* positive category compounds EHA, VPA and PHA, and a negative response to the *in vivo* negative category compound EBA. Six assays responded in this respect: p21, PXR, TCF, ESRE, p53 GENTOX, and anti-PR, identified all these four compounds. When the control compounds (classified as outliers), i.e. DMPA, 4-ene-VPA and PA, being *in vivo* negative, positive, and negative, respectively, are also considered, only one assay, p21, identified all seven compounds correctly in this respect. The other five assays, PXR, TCF, ESRE, p53 GENTOX and anti-PR, all misclassified one of these outliers. PPARa responded to all 9 compounds. Twelve assays didn't respond to any of the 9 compounds. The six statistically best classifying CALUX assays, p21, PXR, TCF, ESRE, p53 GENTOX and anti-PR, show a trend where the potency of the compound increases with chain length (see data matrix): the lowest effect concentration (LEC) for EHA > VPA > PHA (only for ESRE LEC VPA= LEC PHA). Of the thirteen responding assays, none was responding to MHA, similar as the observation for EBA. To the negative control DMPA, only the anti-AR assay responded. Also, to the structurally most closely related MPA, for which no *in vivo* data exist, none of the assays responded.

Table 15. CALUX result table. Results are displayed as LECs in LogM.

compound	Cytotox20%	ERa	ERa-anti	AR	AR-anti	PR	PR-anti	GR	GR-anti	TRb	TRb-anti	RAR	LXR	PXR	PPARa	PPARd	PPARg	AhR	Hif1a	TCF	AP1	ESRE	NFkB	Nrf2	p21	p53 GENTOX
PHA					-3.5		-3.6			-3.5				-4.3	-4.5	-3.0	-3.9	-3.0		-4.3	-4.0	-4.0		-4.0	-3.7	-4.0
EHA					-3.4		-3.0							-3.6	-4.0					-3.5		-3.0		-3.0	-3.0	-3.1
VPA							-3.1			-3.0				-4.1	-4.0					-4.0		-4.0			-3.2	-3.6
4-ene VPA										-3.0				-3.5	-4.1					-3.5		-3.0			-3.1	-3.3
PA										-3.4				-3.5	-3.5					-3.5		-3.0		-3.2		-3.5
DMPA					-3.5										-3.7											
MHA															-4.5											
EBA															-3.5											
MPA															-3.5											

The abbreviations used for the assays are explained in Table 6.

Corresponding mouse and human OED values

Mouse and human OED values for CALUX reporter specified *in vitro* effect values were calculated for the five compounds for which PBPK models were available. The OED values correspond to the indicated *in vitro* specified effect concentrations, i.e. it is the external dose needed to achieve such an internal equipotent target concentration. These are not depicted here, but can be viewed in Annex I. They do not lead to different conclusions as described above.

4.2.2. In silico models

QSAR for prediction in vivo induction of exencephaly in rodents

A more detailed description of the results is provided in Annex II. This model uses both structural characteristics as well as HDAC inhibiting capacity in UKN1 cells of the tested compounds. Based on this QSAR–linear discriminant analysis the known *in vivo* active analogues VPA, EHA, PHA, and 4-ene-VPA have high posterior probabilities for being active and the model, therefore, predicts them as inducers of exencephaly. The other analogues MHA, MPA, PA, and DMPA are predicted with high posterior probabilities as being non-inducers of exencephaly. EBA is predicted to be a weak inducer of exencephaly. The predictions for these negative compounds are in line with their reported *in vivo* observation with respect to exencephaly, save for MPA, which has not been tested for this endpoint.

Toxicokinetic models

Bio-kinetic modelling

The results of the bio-kinetic modelling, i.e. transforming nominal concentrations into embryo-total (ZET/reporter), cell-total (CALUX reporters), and medium unbound (mEST, UKN1) is incorporated into the results description section 4.2.1 for the various experimental models (see above).

PBPK modelling

The results of the reverse PBPK modelling, i.e. transforming the bio-kinetic modelling derived dose descriptors into OEDs is incorporated into the justification section 4.3.2 for the various experimental models (see below).

From this data, i.e. predicted OEDs corresponding to these *in vitro* target concentrations generated for the various specified models, it can be concluded that OED values for humans are in general lower than those for mice. This suggests humans build up higher target concentrations at similar external dose as compared to mice; for the 5 compounds explored here, this differs per compound and per *in vitro* model system.

Concerning the differences among chemicals with regard to effective delivery[5] at the target: for mouse this is predicted to be most efficiently done by VPA, followed by MHA and MPA and EBA, with EHA being least efficient, across all models. The difference between MHA and EHA being a factor of 2 to 3, MHA being more effective. For human predictions differences are smaller and different, EHA and VPA being most efficient, shortly followed by MHA, and subsequently by MPA, and EBA. Here differences between MHA and EHA being around 20 to 40 percent only, with EHA being more effective.

4.2.3. Classification and data analysis models

Dempster-Shafer Theory approach

The Dempster-Shafer theory (DST) (Shafer G., 1976; Dempster AP, 1967) that was applied to the data is an extension of generalised Bayesian statistical inference in which evidence

[5] relative delivery efficacy: ratio of {*in vitro* EC_{10}} and associated {*in vivo* OED} for a chemical.

can be associated with multiple sources. DST represents a rigorous decision-theory approach that provides a framework to generate predictions, estimate the uncertainty associated with each prediction, and combine multiple sources of evidence resulting in a weight-of-evidence (WoE) prediction by quantitatively accounting for the reliability of each of the individual sources (details in Annex III). In general, this decision theory will support the decision making process done by the toxicologist.

In this submission, DST was used to combine the evidence from different *in vitro* assays for the source compounds in order to provide a WoE estimate for the target compound with respect to the *in vivo* neurodevelopmental toxicity outcome. Binary data were used, indicating active/inactive per assay result. The different assay sets are depicted below in the column 'assay set' of Table 16.

First the subset of assays were identified, which gives the best validation results from a leave-one-out (LOO) cross validation. This LOO validation enables to detect the reliability, positive prediction accuracy (PPV) and negative prediction accuracy (NPV). The LOO procedure is a standard approach for selecting a set of sources (in this case assays) based only on the training set (in this case the source compounds).

DST analysis indicated that developmental and neurodevelopmental effects is predicted with 100% certainty from the current *in vitro* assay results.

Using the correlations between *in vivo* and *in vitro* outcomes from the assays for the 7 source compounds the *in vivo* predicted outcome for the target compound in all 3 analysis is that MHA is not an *in vivo* developmental and/or neurodevelopmental toxicant (Table 16; Annex III). The same conclusion was obtained for MPA, though lacking *in vivo* data, but included for its structural similarity to MHA, and for exploring its *in vitro* response profile.

Table 16. Results from DST analysis on target compound prediction.

id	belief	plausibility	classification	classification	true_class	assay set
MHA	0	0	low	non neurodevelopmental	non neurodevelopmental	ZET assay, Pericardial a/o yolk oedema, EC10, Embryo-total ZET assay, Small eyes, EC10, Embryo-total ZET assay, Jitter/tremor, EC10, Embryo-total UKN1, Cell viability, EC10, Medium unbound
MPA	0	0	low	non neurodevelopmental	not tested	ZET assay, Pericardial a/o yolk oedema, EC10, Embryo-total ZET assay, Small eyes, EC10, Embryo-total ZET assay, Jitter/tremor, EC10, Embryo-total UKN1, Cell viability, EC10, Medium unbound
MHA	0	0	low	non developmental and neurodevelopmental	non developmental and neurodevelopmental	ZET assay, All 5 read outs, All EC10, Embryo-total ZET Reporter, CHA cranio-facial deformation, EC10, Embryo-total mEST, Cell viability D3, ID10, Medium unbound UKN1, Cell viability, EC10, Medium unbound CALUX, All 14 assays, LEC, Cell total
MPA	0	0	low	non developmental and neurodevelopmental	not tested	ZET assay, All 5 read outs, All EC10, Embryo-total ZET Reporter, CHA cranio-facial deformation, EC10, Embryo-total mEST, Cell viability D3, ID10, Medium unbound UKN1, Cell viability, EC10, Medium unbound CALUX, All 14 assays, LEC, Cell total
MHA	0	0	low	non developmental and neurodevelopmental	non developmental and neurodevelopmental	CALUX, All 14 assays, LEC, Cell total
MPA	0	0	low	non developmental and neurodevelopmental	not tested	CALUX, All 14 assays, LEC, Cell total

Bayesian Automatic Classification approach

The results of the training of the classifier are given in Annex III.

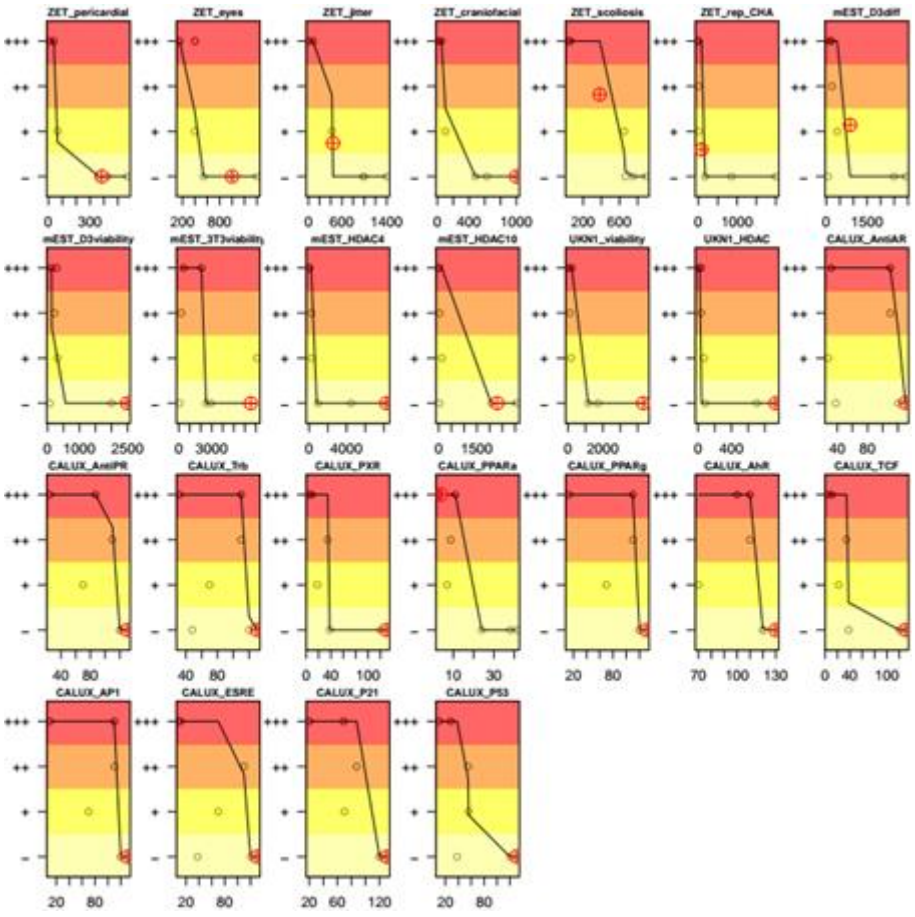
In vitro 10% effect measurements

If only the best tests and relevant to neurodevelopmental toxicity are used (ZET_pericardial, ZET_eyes, ZET_jitter, UKN1_viability), the classifier assigns MHA to category negative with probability 93% of the case and strongly positive with probability 7%. See Figure 7 for the assignment of MHA by those classifiers.

If only the best tests and relevant to neurodevelopmental and developmental toxicity are used (All five ZET read outs, ZET reporter for cranio-facial deformation, mEST cell viability at day 3, UKN1 cell viability and all 14 CALUX assays), the classifier assigns MHA to category negative with probability 90% of the case and strongly positive with probability 10%. See Figure 8 for the assignment of MHA by those classifiers. The strongly positive assignments come from the two indicated tests.

Figure 7. Results of automatic classification of MHA (red crossed circle) given its response in the various assays relevant to the underlying AOP.

The black circles mark the positions of the source chemicals used for training the classifier.



Corresponding *in vivo* OEDs, mouse

If only the best tests and relevant to neurodevelopmental toxicity are used (ZET_pericardial, ZET_eyes, ZET_jitter, UKN1_viability), the classifier assigns MHA to category negative with probability 39%, mildly positive with probability 37%, positive

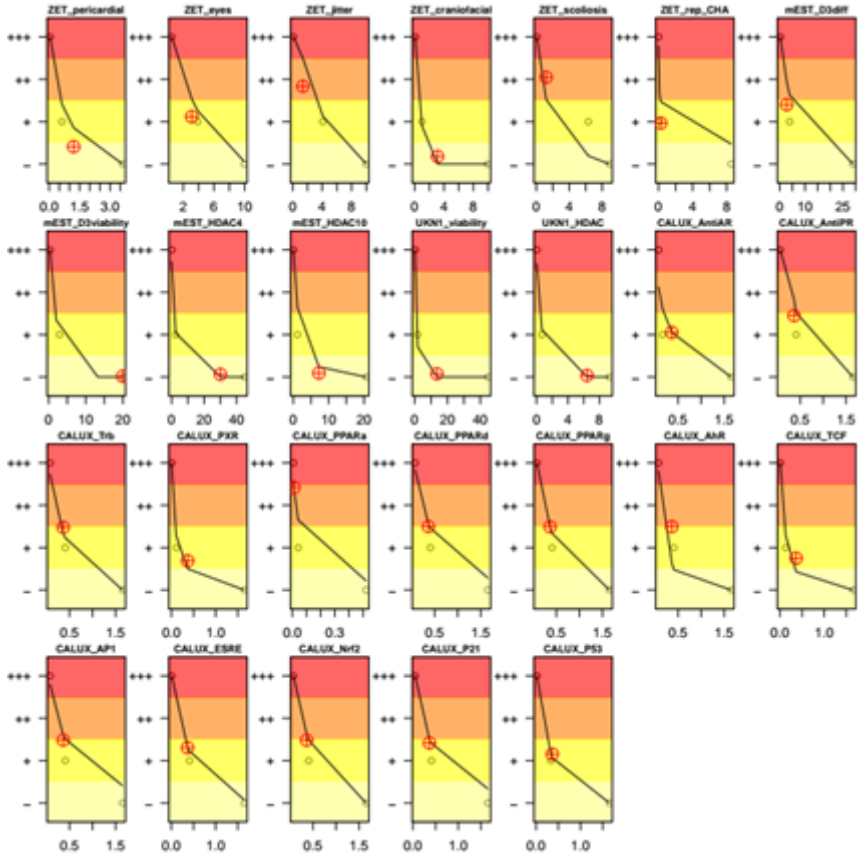
with probability 22%, and strongly positive with probability 1%. See Figure 8 for the assignment of MHA by those classifiers.

If only the best tests and relevant to neurodevelopmental and developmental toxicity are used (All five ZET read outs, ZET reporter for cranio-facial deformation, mEST cell viability at day 3, UKN1 cell viability and all 14 CALUX assays), the classifier assigns MHA to category negative with probability 20%, mildly positive with probability 44%, positive with probability 30% and strongly positive with probability 4%. See Figure 8 for the assignment of MHA by those classifiers.

The low number of training chemicals (only three) is probably responsible for the uncertainty in the predictions. The pharmacokinetic correction also seems to increase the predicted toxicity.

Figure 8. Results of automatic classification of MHA (red crossed circle) given its response in the various assays relevant to the underlying AOP, after pharmacokinetic correction for the mouse.

The black circles mark the positions of the source chemicals used for training the classifier.



Corresponding *in vivo* OEDs, human

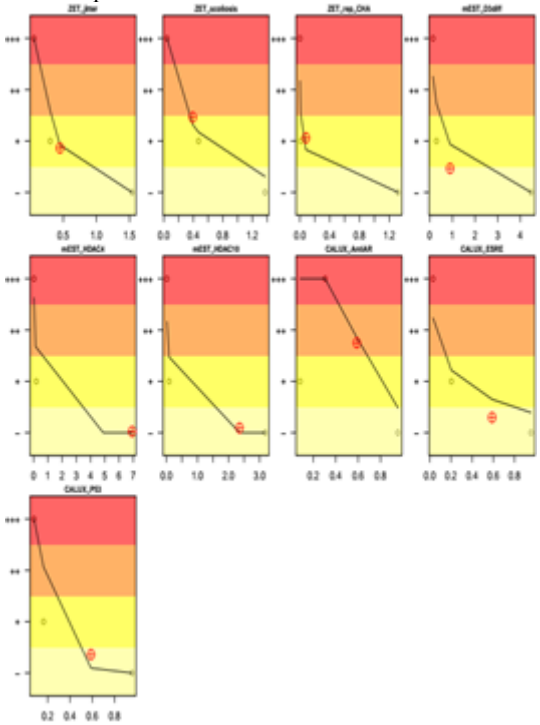
If only the best tests and relevant to neurodevelopmental toxicity are used (ZET_pericardial, ZET_eyes, ZET_jitter, UKN1_viability), the classifier assigns MHA to category negative with probability 69%, positive with probability 18%, and strongly positive with probability 3%. See Figure 9 (next page) for the assignment of MHA by those classifiers.

If only the best tests and relevant to neurodevelopmental and developmental toxicity are used (All five ZET read outs, ZET reporter for cranio-facial deformation, mEST cell viability at day 3, UKN1 cell viability and all 14 CALUX assays), the classifier assigns MHA to category negative with probability 50%, mildly positive with probability 41%, positive with probability 8%, and strongly positive with probability 1%. See Figure 9 (next page) for the assignment of MHA by those classifiers.

The low number of training chemicals (only three) is probably responsible for the uncertainty in the predictions.

Figure 9. Results of automatic classification of MHA (red crossed circle) given its response in the various assays relevant to the underlying AOP, after pharmacokinetic correction for the mouse.

The black circles mark the positions of the source chemicals used for training the classifier.



For MPA, the analogue without *in vivo* data included because of its close structural similarity to the target chemical MHA, also predictions were made to see whether it would fit the trend in the category, i.e. going from *in vivo* positives at the right side (PHA, VPA, and EHA in decreasing potency) to the *in vivo* negatives at the left side (EBA), and to verify whether it would indirectly provide support to any read-across to MHA. Corresponding Figures to the below described analysis results are to be found in Annex III.

Based on EC_{10} values for neurodevelopmental toxicity tests (see above with MHA), the classifier assigns MPA to category negative with probability 99%, and to strongly positive with probability 1%. If all tests relevant to neurodevelopmental and developmental toxicity are used (see above with MHA), the classifier assigns MPA to category negative with probability 96% and to strongly positive with probability 4%.

Based on the correspondingly derived *in vivo* OEDs for mouse for neurodevelopmental toxicity tests (see above with MHA), the classifier assigns MPA to category negative with

probability 90%, mildly positive with probability 10%. If all tests relevant to neurodevelopmental and developmental toxicity are used (see above with MHA), the classifier assigns MPA to category negative with probability 87%, mildly positive with probability 13%.

For the corresponding *in vivo* OEDs for human the classifier assigns MPA to category negative with probability 58%, mildly positive with probability 37%, positive with probability 4%, and strongly positive with probability 1% for neurodevelopmental toxicity tests, and to category negative with probability 79%, mildly positive with probability 19%, positive with probability 2%, and strongly positive with probability 1% for tests relevant to neurodevelopmental and developmental toxicity.

Biological Fingerprint Classification approach

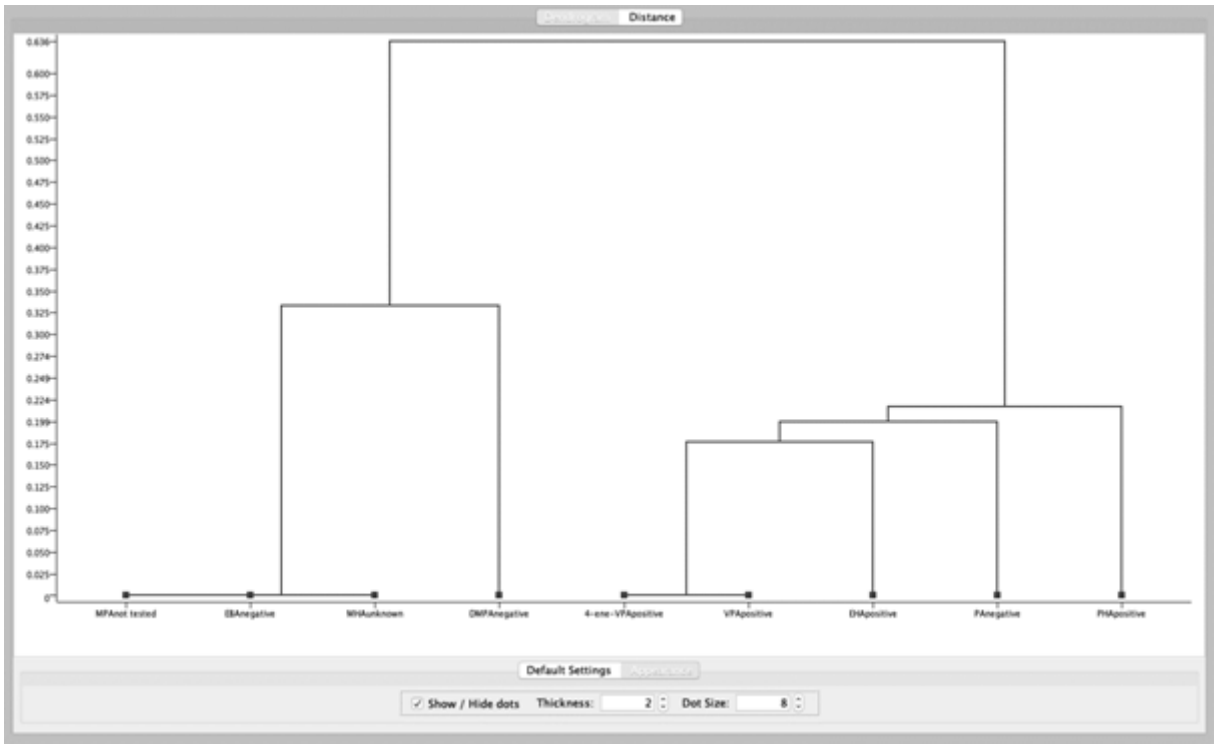
For the analysis, the CALUX battery profile of chemicals was converted into fingerprints that were subsequently assessed on distance between chemicals using the Dice algorithm: Table 17 shows the results of this.

Table 17. Fingerprint distances among chemicals based on CALUX responses.

Key	MHAunknown	PAnegative	DMPAnegative	EBAnegative	MPAnot tested	EHApositive	VPApositive	PHApositive	4-ene-VPApositive
MHAunknown	0.00	0.75	0.33	0.00	0.00	0.80	0.78	0.87	0.78
PAnegative	0.75	0.00	0.78	0.75	0.75	0.25	0.20	0.33	0.20
DMPAnegative	0.33	0.78	0.00	0.33	0.33	0.64	0.80	0.75	0.80
EBAnegative	0.00	0.75	0.33	0.00	0.00	0.80	0.78	0.87	0.78
MPAnot tested	0.00	0.75	0.33	0.00	0.00	0.80	0.78	0.87	0.78
EHApositive	0.80	0.25	0.64	0.80	0.80	0.00	0.18	0.22	0.18
VPApositive	0.78	0.20	0.80	0.78	0.78	0.18	0.00	0.27	0.00
PHApositive	0.87	0.33	0.75	0.87	0.87	0.22	0.27	0.00	0.27
4-ene-VPApositive	0.78	0.20	0.80	0.78	0.78	0.18	0.00	0.27	0.00

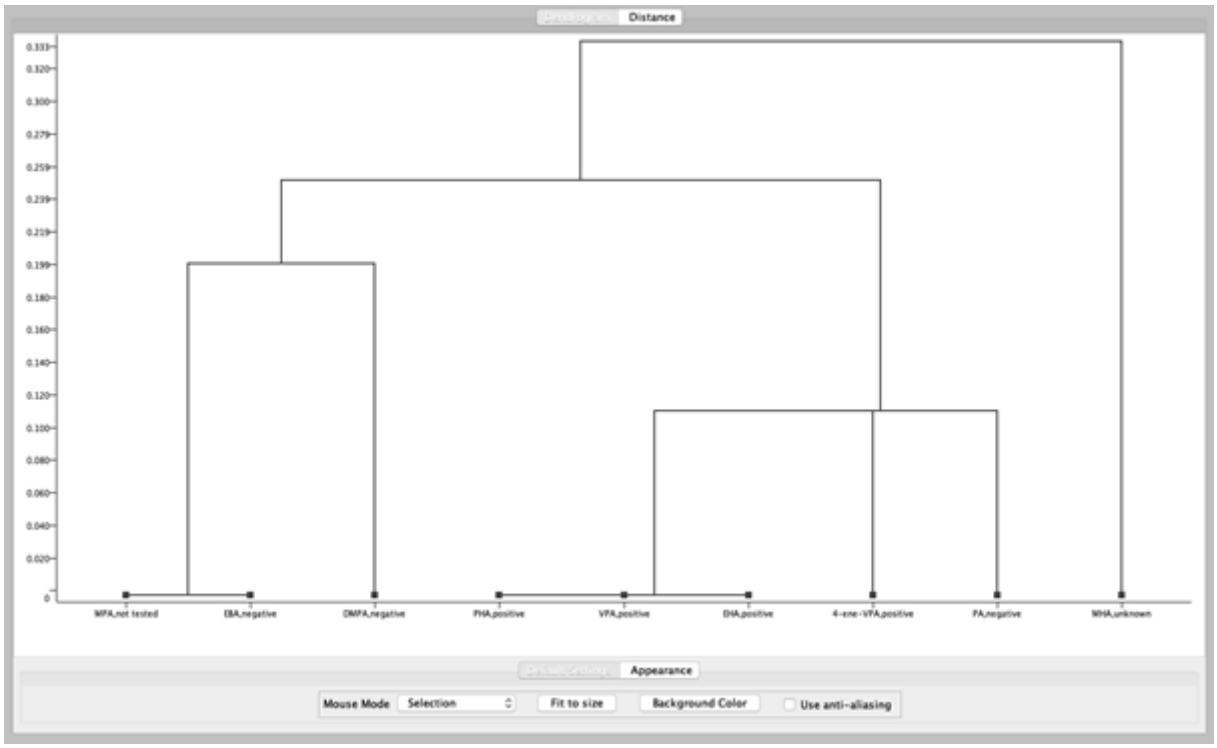
The below dendrogram (Figure 10) made on this basis shows that *in vivo* positive and negative chemicals each cluster together (note PA showing false positive response), and that MHA nicely clusters with the negative compounds,

Figure 10. Dendrogram of distance case study chemicals based on CALUX response fingerprints.



Next this CALUX fingerprint was used as a single entry into an analysis where this CALUX response per chemical was combined with the EC_{10} values for most sensitive endpoint in ZET, i.e. pericardial and/or yolk oedema, the response in ZET reporter (cranio-facial deformation), mEST (differentiation), and UKN1 (viability). The dendrogram that resulted from this analysis is shown in Figure 11.

Figure 11. Dendrogram of distances among chemicals based on battery response fingerprints (with CALUX responses as a single entry).



Interestingly, the clustering of *in vivo* negative and positive chemicals is similar, but MHA has moved to a different position, independent from these two clusters. A similar result is obtained when the analysis is performed with the *in vitro* response for neurodevelopmental and developmental effects combined instead of with the most sensitive effect in ZET, and the other indicated *in vitro* responses. The underlying distance values are given in the table below.

MPA (included because of its close structural similarity to MHA, despite absence of *in vivo* data) clusters with the negative source chemical EBA, and the negative control DMPA in both analyses (i.e. only CALUX, and all assays combined).

Table 18. Fingerprint distances among chemicals based on battery responses (with CALUX responses as a single entry).

Key	MHAunknown	PAnegative	DMPAnegative	EBAnegative	MPAnot tested	EHApositive	VPApositive	PHApositive	4-ene-VPApositive
MHAunknown	0.00	0.33	0.60	0.50	0.50	0.43	0.43	0.43	0.67
PAnegative	0.33	0.00	0.43	0.67	0.67	0.11	0.11	0.11	0.25
DMPAnegative	0.60	0.43	0.00	0.20	0.20	0.25	0.25	0.25	0.43
EBAnegative	0.50	0.67	0.20	0.00	0.00	0.43	0.43	0.43	0.67
MPAnot tested	0.50	0.67	0.20	0.00	0.00	0.43	0.43	0.43	0.67
EHApositive	0.43	0.11	0.25	0.43	0.43	0.00	0.00	0.00	0.11
VPApositive	0.43	0.11	0.25	0.43	0.43	0.00	0.00	0.00	0.11
PHApositive	0.43	0.11	0.25	0.43	0.43	0.00	0.00	0.00	0.11
4-ene-VPApositive	0.67	0.25	0.43	0.67	0.67	0.11	0.11	0.11	0.00

4.3. Justification

In this section, we will discuss and conclude on the results that have been obtained with the above described methods. We will also describe here the specific uncertainties of the biological model. Please, note that as no clear mode of action or even AOP has been

identified for the effects induced by the source chemicals, all available model read outs have been monitored for EC_{10} or comparable dose descriptor values, and have been compared in this study.

In chapter 5. all the data of this chapter and previous chapters 2. to 4. will be integrated to derive an overall conclusion on the data-gap filling for the target chemical MHA, the main objective of this study report.

In the results section 4.2 in fact two dose descriptors have been presented and described: firstly, the specified *in vitro* concentrations that caused a 10% change, either as an increase or decrease, of some effect parameter. By comparing these 10% change concentrations, e.g. EC_{10} values, among the category members, their relative intrinsic potency to this effect-induction is obtained. Please, note that as we have within the category members that induce neurodevelopmental effects *in vivo*, as well as a member that is ineffective in this respect, this relative potency is a critical factor in this read across assessment. Secondly, from these 10% change *in vitro* concentrations OED values are calculated, and by comparing these two dose descriptors one can conclude on the relative delivery efficacy of the category members, i.e. the relative external dose needed to achieve the 10% change concentration at the target. The below sections 4.3.1 and 4.3.2 will discuss intrinsic potency *in vitro* as well as delivery efficacy of external dose, respectively.

4.3.1. In vitro DART battery models

ZET

Uncertainty

Critical aspects with ZET testing are the bioavailability of the tested compounds in the model (i.e. chorion (egg shell) permeation), and herewith their solubility in the medium, as well as the biological variability of egg batches used. The latter is well addressed as the laboratory is well experienced with this ZET model, by applying strict criteria in selecting appropriate fertilised eggs, and performing all experiments in triple. Solubility appears well controlled by the use of a strict solubilisation protocol with DMSO.

Assessment

Of the category chemicals investigated in this study, the aliphatic carboxylic acids PHA, VPA, and EHA, that all induce exencephaly in mice *in vivo*, also appear most potent towards inducing pericardial and/or yolk oedema in ZET, i.e. all showing clearly lower total embryo EC_{10} values as compared to the other investigated analogues MHA, MPA, and EBA, and of which EBA is known to be negative for inducing exencephaly in mice *in vivo* (for MPA the *in vivo* response was not investigated). For 4-ene-VPA, the 'positive control' (i.e. it induces exencephaly in mice *in vivo*), we do not have *in vitro* data, unfortunately. It is noted, that the *in vivo* least potent positive analogue EHA, is also the least potent of the *in vivo* positive chemicals in inducing pericardial oedema in ZET. The two 'negative controls', DMPA, and PA, both have relative high total embryo EC_{10} values. MHA would on the basis of its total embryo EC_{10} value for these ZET read outs group with the lower potent analogues, and on this basis be predicted as negative in the *in vivo* assay. This also holds for the structurally most closely related MPA, for which we do not have *in vivo* data.

A similar picture emerges when looking at the two clear developmental effects, cranio-facial deformation, and scoliosis & lordosis: the *in vivo* positive analogues appear clearly more potent than the *in vivo* negative analogues. It is noted, that the *in vivo* least potent

positive analogue EHA, is also the least potent in inducing cranio-facial deformation in ZET. Again, when looking at the total embryo EC_{10} of MHA for inducing these clear developmental effects, it would be predicted negative *in vivo* on this basis. A similar conclusion would be drawn for MPA (not tested *in vivo*).

Regarding the endpoints most related to neurotoxicity, and herewith assumed as most associated to exencephaly *in vivo*, jitter/tremor and small eyes, the following observation is made: the most potent inducers of exencephaly in mice *in vivo*, PHA, and VPA, are most potent in inducing jitter/tremor and/or small eyes as well. The analogue that is negative *in vivo* in this respect, EBA, fails to induce these effects. This also holds for the 'negative control' PA, while MHA, MPA, and the negative control DMPA are similarly weak in this respect. Thus, MHA appears to categorise with its total embryo EC_{10} for small eyes with EBA, and the 'negative controls' DMPA, and PA, while it categorises with EHA for jitter/tremor. Though MHA is less potent as compared to EHA, it is on this basis, therefore, not possible to classify MHA with certainty for its *in vivo* response as absolutely negative or positive.

Conclusion

On the basis of these ZET responses by the target and source chemicals, as well as the 'negative controls', the overall conclusion favours MHA as not being able to induce developmental effects, and herewith exencephaly, in mice *in vivo*.

ZET Reporter

Uncertainty

Next to the aspects outlined for ZET above, for this reporter assay there are two principal elements of uncertainty associated with the interpretation of the data output. The assay is very specific to craniofacial cartilage structures and relies on accurate orientation of the larvae, and presents a challenge for low potency compounds, due to apparently a specific toxicological effects. This uncertainty is reflected in the dose-response curve fitting, reflected in the R^2 value: low for >0.9, medium for between 0,7 and 0,9, and high for <0,7. Another type of uncertainty concerns the interpretation of data derived from compounds for which no dose-response curve could be fitted, due to potent lethal effects at earlier timepoints, i.e. before 5 dpf when CHA is recorded.

Assessment

EHA, VPA, and PHA are about an order of magnitude more potent as the other category members, including MHA. The *in vivo* positive control 4-ene-VPA has a potency comparable to that of the positive category members, while the potency of the negative controls PA, and DMPA group with those of EBA, MPA, and MHA.

Conclusion

Based on these observations for category and control compounds, MHA is considered to not induce cranio-facial deformation in the ZET, confirming the described observation in the above described classical ZET assay.

mEST

Uncertainty

ES cells are obtained from a murine model, which limits the direct translation to humans. The prediction model is based on the inhibition of cardiomyocyte differentiation (mesodermal differentiation), and may well be limited for identifying compounds that specifically target endo- and ectodermal development. Also, uncertainties of this method are caused due to limited metabolic capacity of the ES cells (leading to false negatives when metabolite is toxicant, or false positives if parent is toxicant), as well as the absence of an *in vivo* physiological distribution system. Solubility was well monitored, while for bioavailability it was assumed that extracellular free concentrations mimic those intracellularly. The laboratory is well experienced with this model (high reproducibility, low variability).

Assessment

The observation that EBA, and the 'negative control' DMPA are unable to inhibit D3 differentiation up to the highest nominal dose of 3000 µM is in line with their *in vivo* negative score for induction of exencephaly. PHA, VPA, EHA, and the 'positive control' 4-ene-VPA, that all do induce exencephaly in mice *in vivo*, are also most potent towards inhibiting differentiation in D3, i.e. their medium unbound ID_{10} are the lowest among the investigated analogues, ranging from 116 to 406 µM, somewhat reflecting their *in vivo* potency (PHA>VPA>EHA). One exception here being the 'negative control' PA, having the lowest medium unbound ID_{10} of all investigated chemicals: 71 µM. Though, intrinsically active, it appears to be metabolised quickly under *in vivo* conditions (Brown N.A, 1987). Our target chemical MHA shows a somewhat in between behaviour: its medium unbound ID_{10} is 875 µM, i.e. in between the *in vivo* negative analogues EBA, and DMPA, and the *in vivo* positive analogues VPA, PHA, and EHA. The medium unbound ID_{10} of MHA is higher than that of EHA, which *in vivo* already is a weak inducer of exencephaly.

Regarding HDAC inhibition properties, the picture emerging is quite clearcut: the *in vivo* positive analogues and the positive controls have at least one order higher potency of inhibiting these histone deacetylase enzymes. MHA groups with the *in vivo* negative compounds. This also holds for its structural most similar analogue MPA.

Conclusion

On the basis of differentiation inhibition potency MHA would at most be classified as a weak inhibitor, if at all; less potent than EHA. When HDAC inhibition potency is integrated into this assessment, i.e. when assuming to play a critical role in developmental toxicity induction by these aliphatic carboxylic acids, then MHA would be classified as negative. One could also conclude though that HDAC inhibition and differentiation inhibition are not strongly linked for this chemical.

UKN1

Uncertainty

This human-derived test model detects compounds that interfere with defined biological processes during a specific time period of embryonic neurodevelopment (neural induction,

patterning, differentiation), but not processes and cell types that occur later in embryonic development, such as neurite outgrowth or neural crest cells or mature neurons. Uncertainties of this method are also caused due to limited metabolic capacity of the cells (leading to false negatives when metabolite is toxicant, or false positives if parent is toxicant), as well as the absence of an *in vivo* physiological distribution system. The laboratory is well experienced with this model (high reproducibility, low variability). The method is not fully evaluated concerning sensitivity and specificity, as there are only few DNT positive compounds identified in humans, mainly based on epidemiological studies.

Assessment

Comparing *in vitro* medium unbound concentrations for EC_{10} viability, for induction of marker genes expression, and for inhibition of rosette formation with *in vivo* exencephaly induction potential for the investigated analogues shows the following picture: those analogues inducing exencephaly *in vivo*, i.e. EHA, VPA,and PHA, and the 'positive control' 4-ene-VPA, have relatively low EC_{10} values of 184, 215, 43, and 121 µM, respectively, show clear inhibition of rosette formation at these concentration levels, and marker gene expression changes at somewhat higher concentrations. PHA, though, behaves somewhat differently in this latter respect: this analogue doesn't induce expression changes of marker genes at levels clearly inhibiting rosette formation. The prediction (UNK1) score does classify EHA, VPA, and 4-ene-VPA as positive, and PHA as weakly positive neurodevelopmental toxicants.

The *in vivo* negative compounds, i.e. the analogue EBA, and the 'negative controls' DMPA, and PA, have all clearly higher medium unbound EC_{10} values, i.e. >4476, 1164, and 1710 µM, respectively, than *in vivo* positives. EBA and DMPA also show inhibition of rosette formation at these higher concentrations, but do not change expression of marker genes, while PA shows both gene expression changes, and rosette inhibition at the EC_{10} viability level. MPA, without *in vivo* data, also had a high medium unbound EC_{10} value of >4319 µM (and a UKN1 score of 6.3), showing the gene expression changes, but no inhibition of rosette formation. It may be seriously doubted whether these high concentrations are achieved *in vivo*.

MHA showed some gene expression changes, and inhibition of rosette formation (giving a UKN1 score of 7.9), but only at a relatively high medium unbound EC_{10} value of >4321 µM.

As MHA is the least potent of all analogues in HDAC inhibition, with an EC_{10} of 2048 µM, this activity evidently is not involved in the observed inhibition of rosette formation (i.e. for this chemical).

Conclusion

The effects of MHA, resembles some of the *in vivo* negative compounds. Also, its relatively low HDAC inhibition potential supports the conclusion that MHA would be expected to be negative for induction of exencephaly *in vivo*. The prediction score at a relatively high concentration supports this conclusion.

CALUX Reporters

Uncertainty

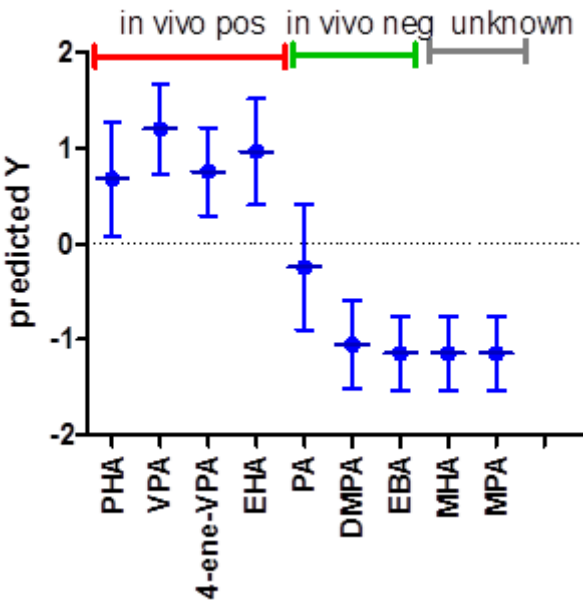
All assays have the same cellular background of U2-OS cells, that have very low expression levels of endogenous receptors, no cross-talk with other receptors, and, therefore, are highly specific and responsive. Uncertainties of this method are also caused due to limited metabolic capacity of the cells (leading to false negatives when metabolite is toxicant, or false positives if parent is toxicant), as well as the absence of an *in vivo* physiological distribution system. If information on bio(in)activation is required, a metabolic module can be added to the CALUX assay to study the activity of metabolites. Solubility was well monitored, while for bioavailability it was assumed that extracellular free concentrations mimic those intracellularly. The laboratory is well experienced with this model (high reproducibility, low variability).

Assessment

When comparing the exencephaly-inducing capability *in vivo* of the test compounds with response induction in CALUX reporter assays, the following picture emerges: the *in vivo* positive compounds EHA, VPA, PHA, and the positive control 4-ene-VPA, are active, i.e. having their LEC at ≤ 1 mM concentration, in 8, 7, 13, and 6 of the 25 assays (leaving out PPARa as non-discriminative here), while for the *in vivo* negative compounds EBA, DMPA, and PA this is 0, 1 and 6, respectively. Additionally, those being *in vivo* most potent, VPA and PHA, also show the lowest LEC values in the CALUX assays: VPA has LEC values ≤ 0.1 mM for 3 assays, for PHA this is for 6 assays, while none of the LEC values for EHA (and the positive control 4-ene-VPA), is ≤ 0.1 mM. MHA doesn't trigger any assay (not counting PPARa, that is triggered by all), and would be grouped with EBA, and DMPA, not with EHA, VPA, PHA, and 4-ene-VPA, which implies it will most probably be negative *in vivo* as well. A similar conclusion can be achieved when considering the chain length of the 6 analogues EBA, MPA, MHA, EHA, VPA, and PHA, and their potential to activate assays: for EBA, MPA, MHA no activations are observed up to 1 mM, while EHA, VPA, and PHA activate 8, 7, and 13 assays up to 1 mM, with VPA, and PHA activating 3, and 6, respectively, even at 0.1 mM. Thus, with increasing chain length, compounds show more potency in activating CALUX assays.

Partial least squares discriminant analysis (PLS-DA) was performed on the CALUX results to build a predictive model for the CALUX panel. A statistical model was built using the analogues that have *in vivo* data, i.e. PHA, VPA, 4-ene-VPA, EHA, PA, DMPA and EBA. Subsequently, MPA and MHA were predicted based on this model. The PLS-DA model was able to differentiate between the *in vivo* positive compounds (+1) and the *in vivo* negative compounds (-1). The CALUX assays identified as most relevant for building this model were ESRE, p21, TCF, PR-anti, PXR, p53 GENTOX; these are the same six assays identified as key factors when considering the 'false positive/negative' rate. The PLS-DA model predicted MHA as negative *in vivo* (Figure 12).

Figure 12. PLS-DA analysis; statistical model was built using PHA, VPA, 4-ene-VPA, EHA, PA, DMPA and EBA. MHA, and MPA were predicted based on this model.



Role of HDAC inhibition on activation profile

VPA is known to inhibit histone deacetylase (HDAC), a negative regulator of gene expression in multiple settings. HDACs are recruited by a variety of transcription factor corepressor complexes and are believed to repress transcription by reducing the level of acetylation of core histones, thereby altering chromatin structure. HDAC inhibition therefore results in an increase in activity of several reporters and cell signalling pathways. The CALUX activity profile of VPA and its analogues could be explained by their activity as HDAC inhibitors, rather than direct interaction with the pathways (Piersma *et al.* 2013).

The CALUX assays that are predominantly activated by the VPA analogues are p21, TCF, ESRE, p53 GENTOX, PXR and anti-PR. All nuclear receptor-mediated transcriptional responses are regulated by HDACs. The reason why not all receptors are being activated by the HDAC inhibitors may be because different HDACs exist with different selectivities. Possibly the valproates inhibit a specific class only. Alternatively, it could be that some background receptor activity may be needed to see an effect of HDAC-inhibition. In our assays, background activity is very low, and possibly with low concentrations of specific ligand the activation may become visible. The four cell signalling pathways p21, p53, TCF (Wnt) and ESRE, on the other hand, have been previously reported to be upregulated by VPA through HDAC inhibition (Phiel *et al.* 2001, Hoti *et al.* 2006, Paradis and Hales Barbara 2015, Segar *et al.* 2017).

HDAC inhibition studies with mEST and UKN1 models (Table 11 and Table 13) have shown that five of the compounds involved in this read-across act as HDAC inhibitors. See Figure 13; all four *in vivo* active compounds VPA, EHA, PHA and 4-ene VPA are identified as HDAC inhibitors, and are active on all four selected CALUX assays. The *in vivo* negative compound PA was already defined as an outlier based on its structure, its

CALUX profile and the PLS-DA analysis; also with respect to HDAC inhibition, it behaves as an outlier, since it is positive here as well.

Figure 13. Relationship between CALUX reporter activation for TCF, ESRE, p21 and p53 and HDAC inhibition properties of case study chemicals.

CALUX: grey cells = not active; values = LEC in Log(M). HDACi: green cells = not active; red cells = active with EC_{50}< 2mM.

compound	TCF	ESRE	p21	p53 GENTOX	HDACi (UKN)
DMPA					2. neg
EBA					2. neg
MPA					2. neg
MHA					2. neg
PA	-3.5	-3.0		-3.5	1. pos
EHA	-3.5	-3.0	-3.0	-3.1	1. pos
4-ene VPA	-3.5	-3.0	-3.1	-3.3	1. pos
VPA	-4.0	-4.0	-3.2	-3.6	1. pos
PHA	-4.3	-4.0	-3.7	-4.0	1. pos

Conclusion

Based on the above statistical reasoning, structure-based analysis, as well as PLS-DA, MHA is predicted as *in vivo* negative for exencephaly induction.

4.3.2. In silico models

QSAR for prediction in vivo induction of exencephaly in rodents

For the target compound MHA of this read-across case study, the model predicts it as a non-inducer of exencephaly. The three *in vivo* positive source compounds, EHA, VPA, and PHA, are predicted as inducers of exencephaly, while a fourth source compound, EBA, known to be negative *in vivo*, is predicted as weak inducer. This span of activity for the source compounds affords all the necessary information to correctly predict the target compound's ability to induce the exencephaly *in vivo*.

Toxicokinetic models

Bio-kinetic modelling

The results of the bio-kinetic modelling, i.e. transforming nominal concentrations into embryo-total (ZET/reporter), medium unbound (mEST, UKN1), and cell-total (CALUX reporters) is incorporated into the results description section 4.2.1 for the various experimental models (see above), and included in the discussion and conclusion section 4.3.1.

PBPK modelling

Differences between chemicals here mainly reside in differences in absorption, plasma-binding and clearance rates. From Annex I it can be deduced that plasma-binding of EHA is higher than that of MHA by a factor of about two (and MPA and EBA equal MHA), while *in vivo* clearance is almost an order lower for EHA (here MPA and EBA double MHA). Although EHA is one of the more lipophilic compounds in the series it is more ionised and so is less rapidly and less completely absorbed when ingested orally, therefore, to account for this higher doses are required to achieve the effective peak concentrations, this is especially pronounced in the mouse model.

The above kinetic aspects are integrated via the reverse dosimetry calculations into OED values. The predicted OED values associated with *in vitro* target concentrations show that EHA is less effective in building these target concentrations in mouse as compared to MHA, by a factor of 2 to 3, while in human on the other hand EHA is slightly more effective in this, by 20-40%. It should be noted that this predicted species difference in toxicokinetics is driven in part, by the use of a more conservative model for the prediction of oral absorption parameters in the human PBPK models (namely the fraction absorbed, fa, and the first-order absorption rate constant, Ka (1/h); a comparison of the models used to predict oral absorption parameters in human and mouse is detailed in Annex III). For arriving at relative *in vivo* toxic potency factors for these chemicals these toxicokinetic differences between MHA and the other category members need to be combined with their toxicodynamic differences, especially the relative potency of MHA towards EBA, and EHA, respectively, is of importance to make a reliable prediction on the *in vivo* (neuro)developmental potency of MHA itself. This comparison is made in section 5.2.

4.3.3. Classification and data analysis models

Dempster-Shafer Theory approach

Using DST that correlates *in vivo* observations (i.e. being neurodevelopmental toxicant or not) and *in vitro* outcomes of the assays for the 7 source compounds, the *in vivo* predicted outcome for the target compound MHA in all 3 analysis (only neurodevelopmental read outs, all read outs, only CALUX read outs) is that it is not neurodevelopmental toxic. For MPA the same conclusion holds.

Bayesian Automatic Classification approach

From this classification analysis, the results for the *in vitro* effect concentrations clearly show MHA to be predicted negative for neurodevelopmental effects induction, as well as induction of developmental and neurodevelopmental effects, with probabilities of 93, and 90%, respectively.

When corresponding human OED values are used as input for this, the negative prediction probabilities for neurodevelopmental effects, and for both developmental and neurodevelopmental effects reduce to 69, and 50%, respectively. For mouse these predictions for neurodevelopmental toxicity turn to 39% for negative, 37% for mild positive, 22% for positive, and 1% for strong positive, while predictions for developmental and neurodevelopmental effects turn to 20% for negative, 44% for mild positive, 30% for positive, and 4% for strong positive.

For MPA, one of the two structurally closest to MHA (next to EHA), for which we also do not have *in vivo* data (but that was included for exploring any category trend), the

corresponding negative prediction probabilities for neurodevelopmental effects, and for both developmental and neurodevelopmental effects, are 99, and 96%, respectively, when based on *in vitro* data. The OED values derived negative prediction probabilities for this analogue for neurodevelopmental effects, and for both developmental and neurodevelopmental effects, being 90, and 87%, for mouse, respectively, and 58, and 79%, for human, respectively.

Biological Fingerprint Classification approach

The results show that based on CALUX responses MHA nicely fits in with the *in vivo* negative compounds. If the CALUX responses are converted into a single response, and combined with responses seen in the other *in vitro* assays, MHA moves away from the negative compounds to an independent position.

5. Strategy for and integrated conclusion of data gap filling

In this chapter, an integrated conclusion is drawn on filling the *in vivo* data-gap of the target chemical MHA on the basis of all results generated with the methodology described in chapter 4. , and on the basis of the approach and assumptions as described in chapters 2. and 3. As all background information as well as all information generated in the context of this study has intrinsic uncertainty that may have bearing on the final conclusion, all this will be outlined in the first section of this chapter to arrive at a final conclusion with regard to the objective of this study in section 2. Therefore, here we will discuss the following aspects: 1) Hypothesis used for the read across; 2) Structural similarity; 3) Similarity of physio-chemical properties; 4) Similarity of toxicokinetics data; 5) Similarity of other supportive data (e.g. data related to key event); 6) Number of analogues used for the read-across; 7) Quality of the endpoint data used for the read-across; 8) Similarity of the endpoint data (among source chemicals); 9) Concordance and weight of evidence of all data used for justifying the hypothesis, and finally 10) Overall uncertainty of the read across.

5.1. Uncertainty

5.1.1. Hypothesis used for the read-across

Source compounds are selected on a firm and transparent structural basis, showing high structural similarity to the target MHA. However, within this well-defined category a remarkable difference in toxicological profile is observed: some show induction of exencephaly in mice (and man), a neurodevelopmental disorder, while others don't. There is yet no clearcut structural feature to be linked with this specific toxicity. To reliably classify, MHA within this category with regard to this neurodevelopmental toxicity, response data within specific *in vitro* models is generated. The *in vivo* data suggest that in this case study structure-activity relationships may be subtle.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is low to medium.

5.1.2. Structural similarity

The structural similarity basis of this read across is quite reliable: all category members are 2-branched aliphatic carboxylic acids. The category includes members differing only one carbon-atom with the MHA molecule, i.e. MPA, and EHA, respectively. Nonetheless, the *in vivo* data suggest that in this case study structure-activity relationships may be subtle.

It is noted that some of the aliphatic carboxylic acids have asymmetric carbon atoms, and thus may exist as racemates: i.e. MPA, MHA, EHA, PHA, and 4-ene-VPA. The *in vivo* data in the NMRI mouse model were generated with the racemates, as far as data on this was retrievable. Also, all *in vitro* tests described in this report have been performed with racemates of these chemicals as well.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is low.

5.1.3. Similarity of physio-chemical properties

(see remark on asymmetric carbon atoms for some aliphatic carboxylic acids under 5.1.2).

The structural similarity basis of this read across is the 2-branched aliphatic carboxylic acid structure, with only minimal chain-length variations. When glancing at the typical physchem properties going from small chain members to longer ones relatively small ranges are observed (rounded values): 116 to 172 g/mol for molecular weight, 15 to 59ºC, and from 196 to 269ºC for melting and boiling points, 1,7 to 3,2 or from 2,0 to 3,9 for experimental or predicted LogPow ranges, 18 to 0,3 g/L for water solubility shows, while experimental pKa constant values appear rather similar: 4,7 to 4.6, leading to the conclusion that these compounds are all non-volatile, well to reasonable water soluble, will not bio-accumulate, and are weak acids of quite comparable strength.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is low.

5.1.4. Similarity of toxicokinetics data

(see remark on asymmetric carbon atoms for some aliphatic carboxylic acids under 5.1.2).

Bio-kinetic modelling

The results of the bio-kinetic modelling, i.e. transforming nominal concentrations into embryo-total (ZET/reporter), medium unbound (mEST, UKN1), and cell-total (CALUX reporters) is performed as described in section 4.2.1. For CALUX reporters calibrated models were used. Due to unavailability of bio-analysis data for ZET, mEST, and UKN1 uncalibrated models were used.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is low to medium.

Metabolism

Using Meteor Nexus biotransformations were predicted for all category analogues (all these compounds were within its domain). Also, a detailed metabolism pattern (first generation) was generated for the two nearest analogues of MHA, being MPA and EBA on one side, and EHA, and VPA on the other side. From this, it is concluded that both biotransformation fingerprints, and concordant metabolite scores are high.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is low to medium.

It is remarked here that all *in vitro* models were metabolically not fully competent (as also indicated at the model descriptions in section 4.3.1). This introduces uncertainties to the dose-response data, and the final conclusions on intrinsic potency.

PBPK modelling

A trend between fraction unbound and side chain length is observed. Only predicted data for plasma protein binding are available (from four different QSAR models). All compounds are within the models' applicability domain. The highest and lowest predict ppb value was used for the *in vitro* to *in vivo* calculation. By this approach a range of human equivalent doses is predicted, which gives in our view a more realistic picture than taking only one single value.

Intrinsic hepatic clearance was measured for all grouped compounds in HepaRG cells. It turned out that the longer chain analogue PHA was outside the applicability domain of

these assays, whereas VPA and the majority of the shorter chain analogues could be measured.

A functional PBPK model could be build based on *in vivo* data for one analogue, VPA. This model predicts bioavailable concentrations of VPA in plasma and liver very well. A good prediction was also obtained by parameterising the PBPK model with *in vitro* values on ppb and intrinsic hepatic clearance. This proof of concept gives confidence in the IVIVE approach used for all analogues in this case study.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is low to medium.

5.1.5. Similarity of other supportive data (e.g. data related to key event)

Sections 3.5, 3.6 and 3.8 describe and discuss supportive biological information showing relatively similar biological behaviour of the category members where it concerns HDAC inhibition as common initial key event for neurodevelopmental members (section 3.5), and identical chemical-biological interaction results for all category members for DNA and protein binding, skin sensitisation, and irritation for eyes and skin using OECD Toolbox profilers.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is low.

5.1.6. Number of analogues used for the read-across

The number of analogues used is limited, but this is due to the fact that constraints set to source chemicals, being very high structural similarity, and having *in vivo* data from the *same* test-protocol.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is low.

5.1.7. Quality of the endpoint data used for the read-across

The *in vivo* data of the source chemicals are discussed in section 3.3, all come from a specific mouse species protocol from one laboratory, excluding variations possibly due to species, laboratory, diet etc. The *in vitro* and *in silico* data used to built and support the biological read across in this case study are described in sections 4.1 to 4.3, all demonstrated as suitable for identification of (neuro)developmental toxicants.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is low to medium.

5.1.8. Similarity of the endpoint data (among source chemicals)

As a general remark: all generated biological data, be it *in vivo* or *in vitro* intrinsically will show variation in responses. Repetition of *in vivo* studies performed according to OECD standards may show differences in potency of about one order of magnitude between studies, and even differences in specific observed toxicological effects (e.g. Janer *et al.*, 2007, 2008).

In this case study *in vivo* data of the source chemicals come from a specific standard protocol developed by Nau and colleagues, all *in vitro* data generated within this case study

to build the biological evidence for read across is performed for source and target chemicals under exactly the same conditions.

Compared to the overall uncertainty of this assessment the uncertainty assigned to this aspect is relatively low.

5.1.9. Concordance and weight of evidence of all data used for justifying the hypothesis

The structural similarity of these 2-branched aliphatic carboxylic acids has been described as one commonality of the category members. Also, phys-chem properties, and chemico-biological reactivity appear quite comparable.

Predictions on metabolism show reasonable similar profiles as well. On the other hand, it should be mentioned that all *in vitro* models are metabolically not fully competent, which introduces an uncertainty with regard to quantitatively using and comparing *in vitro* responses for concluding on *in vivo* responses and deriving OEDs.

The *in vitro* data generated in this case study show quite consistent effects for the category members, both for the target chemical as well as for source chemicals. There are differences in potency of inducing the various monitored effects, though: from this, it is concluded that MHA is most often grouping with the *in vivo* negative analogue EBA, and with the negative controls PA, and DMPA. Although in some cases it shows somewhat higher potency of inducing *in vitro* effects than these *in vivo* negative chemicals, in all cases its effective concentration is higher than that of EHA, mostly about an order of magnitude, if effective at all (i.e. in some models highest tested doses of MHA didn't show any effects). This is also reflected in the HDAC inhibition potential: MHA clusters with *in vivo* negative category member EBA, and the negative controls.

It is noted that HDAC inhibition on the one hand, and differentiation inhibition in mEST cells, as well as inhibition of rosette formation in UKN1 cells on the other hand, do not fully parallel for MHA: suggesting these two activities are not linked for this chemical (in at least these two models).

When applying integrating statistical analysis to this *in vitro* dataset, DST analysis predicts MHA to be no *in vivo* developmental and/or neurodevelopmental toxicant. The Bayesian automatic classifier analysis showed similar results for MHA, and MPA, when applied to the *in vitro* dose descriptors, and the Biological Fingerprint analysis gives a similar conclusion for the CALUX reporter data. When combining this CALUX data with the other *in vitro* responses, MHA seems to cluster on its own, at a distant from the *in vivo* positive and negative analogues.

The prediction model, built on known qualitatively observed *in vivo* structure-activity features of VPA analogues and their exencephaly induction on the one hand, HDAC inhibitory potency, and relevant physical-chemical parameters on the other, predicted MHA to be non-exencephaly inducer.

The Bayesian automatic classifier analysis is also applied to the OED data that were derived from the *in vitro* dose descriptors by reverse PBPK modelling. For man this result in a similar conclusion as derived from the *in vitro* data: MHA clusters more with the *in vivo* negative members, than with EHA. A somewhat different picture emerges for the mouse though: as a result of a predicted higher delivery efficacy[6] for MHA, as compared to EHA,

[6] relative delivery efficacy: ratio of {*in vitro* EC_{10}} and associated {*in vivo* OED} for a chemical.

the potency difference between MHA and EHA observed *in vitro* is partly overcome, reducing the difference in external equipotent dose level for these two chemicals.

This above data is rather concordant with regard to the conclusion: MHA clearly has lower intrinsic toxic potential as compared to EHA, and it clusters frequently with the *in vivo* negative analogue and negative control chemicals. In a few *in vitro* assays, and in a few statistical analysis, it rather stands alone in between EHA, and other *in vivo* positive analogues on one side, and the *in vivo* negative analogue, and control chemicals on the other side.

5.1.10. Overall uncertainty of the read across.

Based on the above considerations the overall uncertainty of this assessment is relatively low to medium.

5.2. Overall integrated conclusion

As no clear mode of action or even AOP has been identified for the neural tube effects induced by the source chemicals, and also the biological applicability domains of the applied models is unclear, all available model read outs have been monitored for EC_{10} or comparable dose descriptor values, and have been compared and analysed in this study.

Looking at the relative intrinsic potency of the target chemical MHA in the various *in vitro* models, it always appeared less potent than EHA, and as compared to VPA, and PHA, while its potency was very often comparable to that of MPA, EBA, and DMPA, and PA, the latter three being negative *in vivo* for neurodevelopmental effects. Of the 22 different read outs monitored in this case study 2 *in vitro* battery (combining two cranio-facial deformation read outs) MHA clusters 18 out of 21 times with EBA, one single time with EHA (mEST differentiation), and was 2 out of 21 times in between EBA and EHA values; for one read out (ZET; scoliosis/lordosis) EBA, MHA and EHA cluster all together.

MHA's relative potency to EHA for neurodevelopmental effects varied from 0.7 in ZET to < 0.05 times in UKN1, and for developmental effects from <0.5 in mEST to <0.1 in ZET. Also, when assuming HDAC inhibition to be a key initial step in induction of neurodevelopmental effects, MHA appeared about twenty times less potent than EHA, while being comparable or less potent as MPA, EBA, and DMPA, these latter two chemicals being *in vivo* negative for neurodevelopmental effects (note that MPA has no *in vivo* data).

Data for MHA in mEST and UKN1 models suggest no link between HDAC and differentiation inhibition for this chemical in these two models.

The DST analyses and the Bayesian automatic classifier analysis of this *in vitro* data both predict MHA to be no *in vivo* (neuro)developmental toxicant. The Biological Fingerprint analysis of the CALUX data supports this; when all *in vitro* data is analysed MHA seems to cluster more on its own, apart from the *in vivo* positive as well as negative analogues.

The prediction model, built on structural features (known qualitatively observed *in vivo* structure-activity features of VPA analogues and their exencephaly induction, and relevant physical-chemical parameters) on the one hand, and a biological feature (HDAC inhibitory potency) on the other, predicted MHA to be no exencephaly inducer.

An important aspect next to this relative intrinsic potency is the relative delivery efficacy of these chemicals as predicted by reverse dosimetry modelling. From comparing

calculated OEDs and associated *in vitro* EC_{10} values, it appears that EHA is less effective in building these target concentrations in mouse as compared to MHA, by about a factor of 2 to 3, while in human on the other hand EHA is slightly more effective in this, i.e. by 20-40%.

By comparing structure and potency across the category, starting from EBA on the left, with the smallest side chains, to PHA on the right, with the longest side chains, the NAM data for MPA suggest it to be *in vivo* negative as well (like EBA), while for MHA the data is less well clearcut: it clearly is in between MPA (and EBA) and EHA, i.e. less potent than EHA. The responses of MPA and MHA thus nicely fit in this category ordering principle of side chain length.

Taken overall, i.e. combining these toxicodynamic observations and toxicokinetic predictions for MHA and EHA on the one hand, and for MHA and MPA, and EBA on the other hand, and the observation that EHA is the least potent developmental toxicant within the category (i.e. compared to VPA, and PHA), the conclusion of these observations is that MHA will *in vivo* be less potent as EHA, i.e. either being negative or, possibly, inducing minimal neurodevelopmental or developmental effects at most. For MPA (included for exploring the response trend across the category) the data more strongly suggest it being *in vivo* negative for neurodevelopmental or developmental effects.

6. Acknowledgements

The case study was part of the EU ToxRisk project (EU Horizon 2020 research grant agreement no. 681002). The core authors of this case study document are Dinant Kroese[1], Andre Wolterbeek[1], Tanja Waldmann[2], Jaffar Kisitu[2], Jonathan Blum[2], Nadine Dreser[2], Claudia McGinnis[3], Manuela Jaklin[3], Stefan Kustermann[3], Katharina Brotzmann[4], Thomas Braunbeck[4], Bjorn Koch[5], Ciaran Fisher[6], Oliver Hatley[6], Frederic Bois[6], Ségolène Siméon[7], Richard MacClenon[8], Paul Walker[8], Tony Long[9], Ulf Norinder[10], Barbara Zdrazil[11], Barbara van Vugt-Lussenburg[12], Bart van der Burg[12].

1 TNO Healthy Living, Utrecht, The Netherlands

2 University of Konstanz, Konstanz, Germany

3 Roche Innovation Center Basel, Basel, Switzerland

4 University of Heidelberg, Heidelberg, Germany

5 Leiden University, Leiden, The Netherlands

6 Certara UK Ltd, Simcyp Division, Sheffield, United Kingdom

7 INERIS, Verneuil-en-Hallate, France

8 Cyprotex, Cheshire, United Kingdom [3]

9 Lhasa Limited, Leeds, United Kingdom

10 MTM Research Centre, Örebro University, Sweden

11 University of Vienna, Vienna, Austria

12 BioDetection Systems, Amsterdam, The Netherlands

7. References

Adler S, Basketter D, Creton S, Pelkonen O, Van Benthem J, Zuang V, *et al.* (2011), Alternative (non-animal) methods for cosmetics testing: current status and future prospects – 2010. *Arch Toxicol*, Vol. 85, pp. 367–485.

Ali S., Champagne D., Spaink H.P., Richardson M.K. (2011), Zebrafish embryos and larvae: A new generation of disease models and drug screens. *Wiley Online Library, Birth Defects Research Part C.*, Vol. 93, pp. 115-133.

Balmer NVMK, Weng B, Zimmer VN, Ivanova SM, Chambers E, Nikolaeva S, Jagtap A, Sachinidis J, Hescheler T, Waldmann and Leist M (2012), Epigenetic changes and disturbed neural development in a human embryonic stem cell-based model relating to the fetal valproate syndrome, *Hum Mol Genet*, Vol. 21, pp. 4104-4114.

Becker R, Ankley G, Edwards S, Kennedy S, Linkov I, Meek B, Sachana M, Segner H, Van Der Burg B, Villeneuve D, Watanabe H, Barton-Maclaren TS. (2015), Increasing Scientific Confidence in Adverse Outcome Pathways, Application of Tailored Bradford Hill Considerations for Evaluating Weight of Evidence. *Regul Toxicol Pharmacol*. Vol. 72, pp. 514-537.

Blaauboer BJ. (2010), Biokinetic modeling and *in vitro-in vivo* extrapolations. *J Toxicol Environ Health B Crit Rev.*, Vol. 13, pp. 242-52.

Bowden CL. (2003), Valproate - Review. *Bipolar Disorders* Vol. 5, pp. 189-202.

Braunbeck T., Lammer E., Leist E., Rudolf M. (2005), Towards an alternative for the acute fish LC50 test in chemical assessment: the zebrafish (*Danio rerio*) embryo toxicity test – an update. *ALTEX*, Vol. 22, pp. 87-102.

Braunbeck, T., Lammer, E. (2006), Detailed review paper "Fish embryo toxicity assays". *UBA report under contract no. 20385422*, 298 pp.

Braunbeck, T., Kais, B., Lammer, E., Otte, J., Schneider, K., Stengel, D., Strecker, R. (2015), The fish embryo test (FET): origin, applications, and future. *Environ. Sci. Pollut. Res.*, Vol. 22, pp. 16247–16261.

Brown N. A. (1987), Teratogenicity of carboxylic acids: distribution studies in whole embryo culture, *In Pharmacokinetics in Teratogenesis,* Vol. II. Edited by H. Nau and W. J. Scott. pp. 153-163. CRC Press, Boca Raton, FL.

Bui LM, Taubeneck MW, Commisso JF, Uriu-Hare JY, Faber WD, Keen CL (1998) Altered zinc metabolism contributes to the developmental toxicity of 2-ethylhexanoic acid, 2-ethylhexanol and valproic acid. *Toxicology*, Vol. 126, pp. 9-21.

Burg van der B., Winter R., Man H-y., Vangenechten C., Weimer M., Berckmans P., Witters H., Van der Linden S. (2010a), Optimization and prevalidation of the *in vitro* AR CALUX method to test androgenic and antiandrogenic activity of compounds, *Reprod Toxicol*, Vol. 30, pp. 18-24.

Burg van der B., Winter R., Weimer M., Berckmans P., Suzuki G., Gijsbers L., Jonas A., Van der Linden S., Witters H., Aarts J., Legler J., Kopp-Schneider A., Bremer S. (2010b), Optimization and prevalidation of the *in vitro* ERα CALUX method to test estrogenic and antiestrogenic activity of compounds, *Reprod Toxicol*, Vol. 30, pp. 73-80.

Burg van der B., Van der Linden S.C., Man H.Y., Winter R., Jonker L., Van Vugt-Lussenburg B., Brouwer A. (2013), A panel of quantitative CALUX® reporter gene assays for reliable high throughput

toxicity screening of chemicals and complex mixtures. *In "High throughput screening methods in toxicity testing" (P. Steinberg, ed)*. John Wiley and Sons, Inc. New York. ISBN 9781118065631 pp. 519-532.

Burg van der B., B. Pieterse, H. Buist, G. Lewin, S. C. van der Linden, H. Y. Man, E. Rorije, A. H. Piersma, I. Mangelsdorf, A. P. Wolterbeek, E. D. Kroese, B. van Vugt-Lussenburg (2015a), "A high throughput screening system for predicting chemically-induced reproductive organ deformities." *Reprod Toxicol*, Vol. 55, pp. 95-103.

Burg van der B., E. B. Wedebye, D. R. Dietrich, J. Jaworska, I. Mangelsdorf, E. Paune, M. Schwarz, A. H. Piersma, E. D. Kroese (2015b), "The ChemScreen project to design a pragmatic alternative approach to predict reproductive toxicity of chemicals." *Reprod Toxicol*, Vol. 55, pp. 114-123.

Choi J. and L. A. Donehower (2014), "p53 in embryonic development: maintaining a fine balance." (1420-682X (Print)), *Cellular and Molecular Life Sciences CMLS*, Vol. 55, pp. 38–47.

Costanza R and Hartung T. (2009), Re-evaluation of animal numbers and costs for *in vivo* tests to accomplish REACH legislation requirements for chemicals – a report by the transatlantic think tank for toxicology. *ALTEX* 26, pp. 187–208.

Courage-Maguire C, Bacon CL, Nau H, and Regan CM (1997), Correlation of *in vitro* anti-proliferative potential with *in vivo* teratogenicity in a series of valproate analogues, *Int. J. Developmental Neuroscience*, Vol. 4, pp. 26-32.

Dearden J. C., *et al.* (2013). "QSPR prediction of physico-chemical properties for REACH." *SAR QSAR Environ Res.* Vol. 24, pp. 279-318.

Dempster AP (1967), Upper and Lower probabilities induced by a multivalued mapping. *Ann.Math.Sat.*, Vol. 38, pp. 325-339.

DB-ALM protocol 197, Database on Alternative Methods to Animal Experimentation, European Union Reference Laboratory for alternatives to animal testing, Joint Research Centre, Ispra, https://ecvam-dbalm.jrc.ec.europa.eu/methods-and-protocols/protocol/1584/automated-calux-reporter-gene-assay-procedure/datasheet

Diav-Citrin O, Shechtman S, Bar-Oz B, Cantrell D, Arnon J, Ornoy A. (2008), Pregnancy outcome after in utero exposure to valproate: evidence of dose relationship in teratogenic effect. *CNS Drugs.*, Vol. 22, pp. 325-34.

Duncan S, Mercho S, Lopes-Cendes I, Seni MH, Benjamin A, Dubeau F *et al.* (2001), Repeated neural tube defects and valproate monotherapy suggest a pharmacogenetic abnormality, *Epilepsia*, Vol. 42, pp. 750-3.

ECHA (2017), Read-Across Assessment Framework (RAAF) Reference: ECHA-17-R-01-EN, Cat.nr.ED-02-17-140-EN-N, ISBN: 978-92-9495-758-0 DoI: 10.2823/619212; March 2017, European Chemicals Agency, Helsinki.

Eikel D., Lampen A., and Nau H. (2006) Teratogenic effects mediated by inhibition of histone deacetylase: evidence from quantitative structure activity relationships of 20 valproic acid derivatives, *Chem.Res.Toxicol*, Vol. 19, pp. 272-278.

Embry M.R., Belanger S.E., Braunbeck T., Galay-Burgos M., Halder M., Hinton D.E., Léonard M.A., Lillicrap A., Norberg-King T., Whale G. (2010), The fish embryo toxicity test as an animal alternative method in hazard and risk assessment and scientific research, *Aquat. Toxicol*, Vol. 97, pp. 79-87.

English JC, Deisinger PJ and Guest D. (1998), Metabolism of 2-Ethylhexanoic Acid Administered Orally or Dermally to the Female Fischer 344 Rat. *Xenobiotica*, Vol. 28, pp. 699-714.

European Commission (2006), Regulation No 1907/2006 of the European Par-liament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorization and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives.

Fisher C., Siméon S., Jamei M., Gardner I., Bois Y.F. (2019), VIVD: Virtual *in vitro* distribution model for the mechanistic prediction of intracellular concentrations of chemicals in *in vitro* toxicity assays, Toxicol In Vitro, Vol. 58, pp. 42-50.

Frisch C, Hüsch K, Angenstein F, Kudin A, Kunz W, Elger CE *et al.* (2009), Dose-dependent memory effects and cerebral volume changes after in utero exposure to valproate in the rat. *Epilepsia*, Vol. 50, pp. 1432-41.

Genschow E., Spielmann H., Scholz G., Pohl I., Seiler A. Clemann N., Bremer S. Becker K. (2004), Validation of the embryonic stem cell test in the International ECVAM Validation Study on three *in vitro* embryotoxicity tests. *Alternatives to laboratory animals*, Vol.32, No. 3, pp. 209-44.

Genschow E., Spielmann H., Scholz G., Seiler A., Brown N., Piersma A., Brady M., Clemann N., Huuskonen H., Paillard F., Bremer S., Becker K., (2002), The ECVAM international validation study on *in vitro* embryotoxicity tests: Results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods, *Alternatives to laboratory animals,* Vol.30, No. 2, pp. 151-76.

Georgoff P.E., Nikolian V.C., Bonham T., Pai M.P., Tafatia C., Halaweish I., To K., Watcharotone K., Parameswaran A., Luo R., Sun D., Alam H.B. (2018), Safety and Tolerability of Intravenous Valproic Acid in Healthy Subjects: A Phase I Dose-Escalation Trial, *Clin Pharmacokinet,* Vol. 57, pp. 209-219.

Ghodke-Puranik Y, Thorn CF, Lamba JK, Leeder JS, Song W, Birnbaum AK, Altman RB and Klein TE. (2013), Valproic Acid Pathway: Pharmacokinetics and Pharmacodynamics. *Pharmacogenetics and Genomics*, Vol. 23, pp. 236-241.

Gijsbers L., Van Eekelen H., De Haan L., Swier JM., Heijnk N., Kloet S., Bovy A., Keijer J., Aarts J.M., Van der Burg B., Rietjens I.M. (2013), Induction of peroxisome proliferator-activated receptor γ (PPARγ)-mediated gene expression by tomato (*Solanum lycopersicum L.*) extracts. *J Agric Food Chem.*, Vol. 6, pp. 3419-27.

Hammond, C.L., Schulte-Merker, S. (2009) Two populations of endochondral osteoblasts with differential sensitivity to Hedgehog signalling, *Development,* Vol. 136, No.23, pp. 3991–4000.

Hanahan D., Weinberg RA. (2011), Hallmarks of cancer: the next generation, *Cell*, Vol. 144, pp. 646-74

Hendrickx AG, Nau H, Binkerd P, Rowland J. (1988), Valproic acid developmental toxicity and pharmacokinetics in the Rhesus Monkey: An interspecies comparison, *Teratology*, Vol. 38, pp. 329-45.

Hollert H., Keiter S., König N., Rudolf M., Ulrich M., Braunbeck T. (2003), A new sediment contact assay to assess particle-bound pollutants using zebrafish (*Danio rerio*) embryos, *J. Soils & Sed.* Vol. 3, pp. 197 – 207.

Hoti N., Chowdhury W., Hsieh J.T., Sachs M.D., Lupold S.E., Rodriguez R. (2006), Valproic acid, a histone deacetylase inhibitor, is an antagonist for oncolytic adenoviral gene therapy, *Mol Ther*, Vol. 14, pp. 768-778.

Jamei M, Marciniak S, Edwards D, Wragg K, Feng K, Barnett A, Rostami-Hodjegan (2013), The simcyp population based simulator: architecture, implementation, and quality assurance. *In Silico Pharmacol*, Vol. 1, pp. 9.

Janer G, Hakkert BC, Vermeire T, Piersma AH (2007), A retrospective analysis of the added value of the rat two-generation reproductive toxicity study versus the rat subchronic toxicity study, *Reproductive toxicology*, Vol. 24, pp. 103-13.

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*, Vol. 50, pp. 206-17.

Jentink J, Loane MA, Dolk H, Barisic I, Garne E, Morris JK *et al.* (2010), Valproic acid monotherapy in pregnancy and major congenital malformations, *New Engl.J.Med*, Vol. 362, pp. 2185-93.

Kais B., Schneider K.E., Keiter S., Henn K., Achermann C., Braunbeck T. (2013), DMSO modifies the permeability of the zebrafish (*Danio rerio*) chorion-implications for the fish embryo test (FET), *Aquat. Toxicol,* Vo. 140-141, pp. 229-38.

Kantola-Sorsa E, Gaily E, Isoaho M, and Korkman M (2007), Neuropsychological outcomes in children of mothers with epilepsy. *Journal of the Int. Neuropsychological Society*, Vol. 13, pp. 642-652.

Kari G., Rodeck U., Dicker AP. (2007), Zebrafish : An Emerging Model System for Human Disease and Drug Discovery, *Clinical Pharmacol. & Therapeutics*, Vol. 82, pp. 70-80.

Kilford P.J., Gertz M., Houston J.B., Galetin A. (2008), Hepatocellular binding of drugs: correction for unbound fraction in hepatocyte incubations using microsomal binding or drug lipophilicity data, *Drug Metab Dispos*, Vol. 36, pp. 1194-1197.

Koren G., Nava-Ocampo AA, Moretti ME, Sussman R, and Nulman I. (2006), Major malformations with valproic acid, *Can Fam Physician*, Vol. 52, pp. 441–447.

Krug AK, *et al.* (2013), Human embryonic stem cell-derived test systems for developmental neurotoxicity: a transcriptomics approach, *Archives of Toxicology,* Vol. 87, pp. 123–143.

Lammer E., Carr G.J., Wendler K., Rawlings J.M., Belanger S.E., Braunbeck T. (2009), Is the fish embryo toxicity test (FET) with the zebrafish (*Danio rerio*) a potential alternative for the fish acute toxicity test?, *Comp. Biochem. Physiol*, Vol. 149C, pp. 196-209.

Leist M, Hasigawa N, Rovida C, Daneshiam M, Basketter D, Kiber I, *et al.* (2014), Consensus report on the future of animal-free systemic toxicity testing, *ALTEX*, Vol. 31, pp. 341–56.

Lewin G, Escher SE, Van der Burg B, Simetska N, Mangelsdorf I (2015), Structural features of endocrine active chemicals – A comparison of *in vivo* and *in vitro* data, *Reprod. Toxicol,* Vol. 55, pp. 81-94.

Linden van der SC, von Bergh A, Van Vugt-Lussenburg B, Jonker L, Brouwer A, Teunis M, Krul C and Van der Burg B. (2014), Development of a panel of high throughput reporter gene assays to detect genotoxicity and oxidative stress, *Mutation Res*, Vol. 760, pp. 23-32.

Lloyd KA (2013), A scientific review: mechanisms of valproate-mediated teratogenesis, *Bioscience Horizons*, Vol. 6, pp. 1-10.

Martinez C.S., Feas D.A., Siri M., Igartúa D.E., Chiaramoni N.S., del.V. Alonso S., Prieto M.J. (2018), *In vivo* study of teratogenic and anticonvulsant effects of antiepileptic drugs in zebrafish embryo and larvae, *Neurotox. And Teratox*, Vol. 66, pp. 17-24.

Muñoz-Espín D., M. Cañamero, A. Maraver, G. Gómez-López, J. Contreras, S. Murillo-Cuesta, A. Rodríguez-Baeza, I. Varela-Nieto, J. Ruberte, M. Collado, M. Serrano (2013), "Programmed Cell Senescence during Mammalian Embryonic Development." *Cell*, Vol. 155, No.5, pp. 1104-1118.

Murko C, Lagger S, Steiner M, Seiser C, Schoefer C, and Pusch O (2013), Histone deacetylase inhibitor trichostatin A induces neural tube defects and promotes neural cvrest specification in the chicken neural tube. *Differentiation*, Vol. 85, pp. 55-66.

Murko C, Lagger S, Steiner M, Seiser C, Schoefer C, and Pusch O (2010), Expression of class I histone deacetylases during chick and mouse development. *Int. J. Dev. Biology*, Vol. 54, pp.1527–1537.

Nagel R. (2002), DarT: The embryo test with the zebrafish *Danio rerio* - a general model in ecotoxicology and toxicology, *ALTEX,* Vol. 19, Suppl. 1, pp. 38-48.

Narotsky MG, Kavlock EZ. (1994), Developmental toxicity and structure–activity rela-tionships of aliphatic acids, including dose–response assessment of valproic acid in mice and rats, *Fundam Appl Toxicol*, Vol. 22, pp. 251–65.

Nau H., Zierer R., Spielmann H., Neubert D., Gansau C. (1981), a new model for embryotoxicity testing: teratogenicity and pharmacokinetics of valproic acid following constant rate administration in the mouse using human therapeutic drug and metabolite concentrations, *Life Sci.*, Vol. 29, pp. 2803-2814.

Nau H., Löscher W. (1986), Pharmacologic evaluation of various metabolites and analogs of valproic acid: teratogenic potencies in mice, *Fundamental and applied Toxicology,* Vol. 6, pp. 669-676.

Nau H, Hawk R-S, Ehlers K (1991), Valproic Acid-Induced Neural Tube Defects in Mouse and Human: Aspects of Chirality, Alternative Drug development, Pharmacokinetics and Possible Mechanisms, *Pharmacology & Toxicology*, Vol. 69, pp. 310-321.

NRC (2000), Scientific Frontiers in Developmental Toxicology and Risk Assessment, NAS Press, National Research Council Washington, D.C. USA.

Nostrand, van, J. L., M. E. Bowen, H. Vogel, M. Barna, L. D. Attardi (2017), "The p53 family members have distinct roles during mammalian embryonic development.", *Cell Death Differ.*, Vol. 4. Pp. 575-579. (1476-5403 (Electronic)).

OECD (2009) Test guideline No.455. Stably transfected human estrogen receptor-transcriptional activation assay for detection of estrogenic agonist-activity of chemicals, OECD Publishing, Paris, https://doi.org/10.1787/9789264076372-en.

OECD (2013a) SPSF by the Netherlands: U2-OS cells Transcriptional ERalpha CALUX®- assay for the detection of estrogenic and antiestrogenic chemicals for inclusion in TG455/TG457, OECD, Paris.

OECD (2013b) SPSF by the European Commission: Performance-Based Test Guideline on Androgen Receptor Transactivation Assays (AR CALUX), OECD, Paris

Ogungbenro, K., Aarons, L., Cresim, Epi, C.P.G. (2014), A physiologically based pharmacokinetic model for Valproic acid in adults and children, *Eur J Pharm Sci*, Vol. 63, pp. 45-52.

Ong LL, Schardein JL, Petrere JA, Sakowski R, Jordan H, Humphrey RR *et al.* (1983), Teratogenesis of calcium valproate in rats, *Fundamental & Applied Toxicology*, Vol. 3, pp. 121-6.

Ornoy A. (2009), Valproic acid in pregnancy: how much are we endangering the embryo and fetus, *Reproductive Toxicology*, Vol. 28, pp. 1-10.

Padmanabhan R, Ahmed I. (1996), Sodium valproate augments spontaneous neural tube defects and axial skeletal malformations in TO mouse fetuses [corrected].[Erratum appears in Reprod Toxicol 1996 Nov-Dec;10(6):VI], *Reproductive Toxicology*, Vol. 10, pp. 345-63.

Paradis F.-H. and Hales Barbara F. (2015), Valproic Acid Induces the Hyperacetylation of P53, Expression of P53 Target Genes, and Markers of the Intrinsic Apoptotic Pathway in Midorganogenesis Murine Limbs Birth Defects Research Part B: *Developmental and Reproductive*, Vol. 104, pp. 177-226.

Park SJ, Ogunseitan OA, Lejano RP (2013), Dempster-Shafer Theory Applied to Regulatory Decision Process for Selecting Safer Alternatives to Toxic Chemicals in Consumer Products, *Integrated Environmental Assessment and Management*, Vol. 10, pp. 12–21.

Paulson RB, Sucheston ME, Hayes TG, Paulson GW (1985), Teratogenic effects of valproate in the CD-1 mouse fetus, *Archives of Neurology*, Vol. 42, pp. 980-3.

Pelka, K.E., Henn, K., Keck, A., Sapel B., Braunbeck, T. (2017), Size does matter – Determination of the critical molecular size for the uptake of chemicals across the chorion of zebrafish (*Danio rerio*) embryos, *Aquat. Toxicol*, Vol. 185, pp. 1-10.

Pennanen S, Auriola S, Manninen A, Komulainen H. (1991), Identification of the Main Metabolites of 2-Ethylhexanoic Acid in Rat Urine Using Gas Chromatography-Mass Spectrometry, *Journal of Chromatography B: Biomedical Sciences and Applications,* Vol. 568, pp. 125-134.

Pennanen S, Tuovinen K, Huuskonen H, Komulainen H. (1992), The Development Toxicity of 2-Ethylhexanoic Acid in Wistar Rats, *Toxicological Sciences*, Vol. 19, pp. 505–511.

Pennanen S, Kojo A, Pasanen M, Liesivuori J, Juvonen RO, Komulainen H. (1996), CYP Enzymes Catalyze the Formation of a Terminal Olefin from 2-Ethylhexanoic Acid in Rat and Human Liver, *Human and Experimental Toxicology*, Vol. 15, pp. 435-442.

Péry, Alexandre R. R., James Devillers, Céline Brochot, Enrico Mombelli, Olivier Palluel, Benjamin Piccini, François Brion, Rémy Beaudouin. (2014), "A Physiologically Based Toxicokinetic Model for the Zebrafish *Danio Rerio*." *Environmental Science & Technology*, Vol. 48, No. 1, pp. 781–90. https://doi.org/10.1021/es404301q.

Petrere JA, Anderson JA, Sakowski R, Fitzgerald JE, De La Igelsia FA (1986), Teratogenesis of calcium valproate in rabbits, *Teratology*, Vol. 34, pp. 263-9.

Phiel C.J., Zhang F., Huang E.Y., Guenther M.G., Lazar M.A., Klein P.S. (2001), Histone Deacetylase is a direct target of valproic acid, a potent anticonvulsant, mood stabilizer, and teratogen, *J Biol Chem,* Vol. 276, pp. 36734-36741.

Piersma A. H., S. Bosgra, M. B. van Duursen, S. A. Hermsen, L. R. Jonker, E. D. Kroese, S. C. van der Linden, H. Man, M. J. Roelofs, S. H. Schulpen, M. Schwarz, F. Uibel, B. M. van Vugt-Lussenburg, J. Westerhout, A. P. Wolterbeek, B. van der Burg (2013), "Evaluation of an alternative *in vitro* test battery for detecting reproductive toxicants.", *Reprod Toxicol*, Vol. 38, pp. 53-64.

Riebeling, Christian *et al.* (2011), "The Embryonic Stem Cell Test as Tool to Assess Structure-Dependent Teratogenicity: The Case of Valproic Acid, *Toxicological sciences: an official journal of the Society of Toxicology*, Vol. 120, No.2, pp. 360-370.

Segar K.P., Chandrawanshi V., Mehra S. (2017), Activation of unfolded protein response pathway is important for valproic acid mediated increase in immunoglobulin G productivity in recombinant Chinese hamster ovary cells, *J. Biosci Bioeng*, Vol. 124, pp. 459-468.

Silva MFB, Aires CCP, Luis PBM, Ruiter JPN, Ijlst L, Duran M, Wanders RJA, Tavares de Almeida I. (2008), Valproic Acid Metabolism and its Effects on Mitochondrial Fatty Acid Oxidation: A Review, *Journal of Inherited Metabolic Disease*, Vol. 31, pp. 205-216.

Sipes NS, Martin MT, Reif DM, Kleinstreuer NC, Judson RS, Singh AV, *et al.* (2011) Predictive models of prenatal developmental toxicity from ToxCast high throughput screening data, *Toxicol Sci*, Vol. 124, pp. 109–27.

Schenk B, Weimer M, Bremer S, van der Burg B, Cortvrindt R, Freyberger A, *et al.* (2010) The ReProTect Feasibility Study, a novel comprehensive *in vitro* approach to detect reproductive toxicants, *Reprod Toxicol*, Vol. 30, pp. 200–18.

Scholz G, Genschow E, Pohl I, Bremer S, Paparella M, Raabe H, Southee J, and Spielmann H (1999), Prevalidation of the Embryonic Stem Cell Test (EST) - A new *in vitro* embryotoxicity test, *Toxicology in vitro*, Vol.13, pp. 675-681.

Shafer G (1976) A Mathematical Theory of Evidence, Princeton University Press, ISBN 9780691100425, 314 pp.

Sonoda T, Ohdo S, Ohba K, Okishima T, Hayakawa K (1990), Teratogenic effects of sodium valproate in the Jcl: ICR mouse fetus, *Acta Paediatrica Japonica*, Vol. 32, pp. 502-7.

Sonneveld E, Jansen HJ, Riteco JAC, Brouwer A, Van der Burg B. (2005), Development of androgen- and estrogen-responsive bioassays, members of a panel of human cell line-based highly selective steroid responsive bioassays, *Toxicological Sciences,* Vol. 83, pp. 136–48.

Sonneveld, E., Riteco, J.A.C., Jansen, H.J., Pieterse, B., Brouwer, A., Schoonen, W.G., Van der Burg, B. (2006), Comparison of *in vitro* and *in vivo* screening models for androgenic and estrogenic activities, *Toxicol. Sci.*, Vol. 89, pp. 173-87.

Sonneveld, E., Pieterse, B., Schoonen, W., Van der Burg, B. (2011), Validation of *in vitro* screening models for progestagenic activities: inter-assay comparison and correlation with *in vivo* activity in rabbits, *Toxicology in vitro*, Vol. 25, pp. 545-554.

Strähle, U., Scholz, S., Geisler, R., Greiner, P., Hollert, H., Rastegar, S., Schumacher, A., Selderslaghs, I., Weiss, C., Witters, H., Braunbeck, T. (2012), Zebrafish embryos as an alternative to animal experiments – A commentary on the definition of the onset of protected life stages in animal welfare regulations, *Repro. Tox*, Vol. 33, pp. 128-132.

Thomas SV, Ajaykumar B, Sindhu K, Francis E, Namboodiri N, Sivasankaran S *et al.* (2008), Cardiac malformations are increased in infants of mothers with epilepsy, *Pediatric Cardiology*, Vol. 29, pp. 604-8.

Tomson T, Battino D. (2009), Teratogenic effects of antiepileptic medications - Review. *Neurologic Clinics*, Vol. 27, pp. 993-1002.

Tomson T. (2005), Gender aspects of pharmacokinetics of new and old AEDs: pregnancy and breast-feeding, *Therapeutic Drug Monitoring*, Vol. 27, pp. 718-21.

Turner S, Sucheston ME, De Philip RM, Paulson RB (1990), Teratogenic effects on the neuroepithelium of the CD-1 mouse embryo exposed in utero to sodium valproate, *Teratology*, Vol. 41, pp. 421-42.

Vugt-Lussenburg, van, BMA, Van der Lee RB, Man HY, Middelhof I, Brouwer A, Besselink H, Van der Burg B. (2018), Incorporation of metabolic enzymes to improve predictivity of reporter gene assay results for estrogenic and anti-androgenic activity, *Reproductive Toxicol.*, Vol. 75, pp. 40-48.

Wagner, C. K. (2008), "Progesterone Receptors and Neural Development: A Gap between Bench and Bedside?", *Endocrinology* Vol. 149, No. 6, pp. 2743-2749.

Walker V and Mills GA (2001) Urine 4-Heptanone: a beta-Oxidation Product of 2-Ethylhexanoic Acid from Plasticisers, *Clinica Chimica Acta*, Vol. 306, pp. 51-61.

Waldmann *et al.* (2014), Design Principles of Concentration-Dependent Transcriptome Deviations in Drug-Exposed Differentiating Stem Cells, *Chem Res Toxicol*, Vol. 27, pp. 408-420.

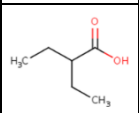
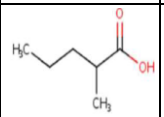
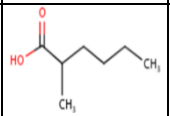
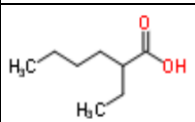
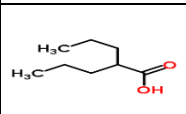
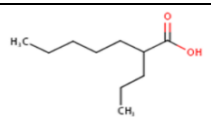
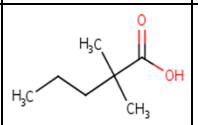
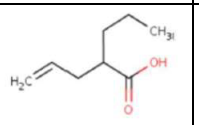
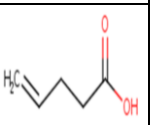
Wetendorf, M. and F. J. DeMayo (2012), "The progesterone receptor regulates implantation, decidualization, and glandular development via a complex paracrine signaling network.", *Mol Cell Endocrinol*, Vol. 357, No. 1-2, pp. 108-18. (1872-8057 (Electronic)).

Wiltse J (2005), Mode of action: inhibition of histone deacetylase, altering WNT-dependent gene expression, and regulation of beta-catenin - developmental effects of valproic acid, *Critical Reviews in Toxicology*, Vol. 35, pp. 727-738.

Winiwarter, S., Ax, F., Lennernas, H., Hallberg, A., Pettersson, C., Karlen, A. (2003), Hydrogen bonding descriptors in the prediction of human *in vivo* intestinal permeability, *J Mol Graph Model*, Vol. 21, pp. 273-287.

Winiwarter, S., Bonham, N.M., Ax, F., Hallberg, A., Lennernas, H., Karlen, A. (1998), Correlation of human jejunal permeability (*in vivo*) of drugs with experimentally and theoretically derived parameters. A multivariate data analysis approach, *J Med Chem*, Vol. 41, pp. 4939-4949.

Annex I. Data-matrix

	Source and Target Compounds								
	Source 1	Source 2	Target	Source 3	Source 4	Source 5	Outlier 1	Outlier 2	Outlier 3
CAS	88-09-5	97-61-0	4536-23-6	149-57-5	99-66-1	31080-39-4	1185-39-3	1575-72-0	109-52-4
Name	2-ethyl butyric acid	2-methyl pentanoic acid	2-methyl hexanoic acid	2-ethyl hexanoic acid	2-propyl pentanoic acid	2-propyl heptanoic acid	2-dimethyl pentanoic acid	2-propyl pentenoic acid	pentenoic acid
Abbreviation	EBA	MPA	MHA	EHA	VPA	PHA	DMPA	4-ene-VPA	PA
Structure									
Structure (smiles)	CCC(CC)C(=O)O	CCCC(C)C(=O)O	CCCCC(C)C(=O)O	CCCCC(CC)C(=O)O	CCCC(CCC)C(=O)O	CCCCCC(CCC)C(=O)O	CCCC(C)(C)C(=O)O	CCCC(CC=C)C(O)=O	C=CCCC(=O)O
Chain length from position 2	2 / 2 / 0	3 / 1 / 0	4 / 1 / 0	4 / 2 / 0	3 / 3 / 0	5 / 3 / 0	3 / 1 / 1	3 / 3 / 0	3 / 0 / 0
	Structural similarity rel. to 4536-23-6								
RDKit Fingerprint	59	83	100	82	69	77	–	–	–
FCFP4 Fingerprint	55	83	100	68	62	52	–	–	–
Layered Fingerprint	60	84	100	86	78	81	–	–	–
RDKit Descriptors	72	76	100	76	75	28	–	–	–
	Summary of Data Gap Filling								
	Source 1	Source 2	Target	Source 3	Source 4	Source 5	Outlier 1	Outlier 2	Outlier 3
CAS	88-09-5	97-61-0	4536-23-6	149-57-5	99-66-1	31080-39-4	1185-39-3	1575-72-0	109-52-4
Name	2-ethyl butyric acid	2-methyl pentanoic acid	2-methyl hexanoic acid	2-ethyl hexanoic acid	2-propyl pentanoic acid	2-propyl heptanoic acid	2-dimethyl pentanoic acid	2-propyl pentenoic acid	pentenoic acid

Target endpoint 1 (neuro-) developmental toxicity in rodent studies Experimental result: NOAEL rats & mice		No studies available	No studies available		Rats: neurodev.effects at 500 mg/kg bw; dev. effects at 250 mg/kg bw.	Mice: (neuro)dev.effects at 200 mg/kg bw Rats: neurodev.effects at 720 mg/kg bw; dev. effects at 100 mg/kg bw.	No studies available	No studies available	No studies available	No studies available
Integrated conclusion				derived result						
Target endpoint 2 exencephaly-induction in NMRI mice (Nau *et al.*, 1986) Experimental result: Potency relative to VPA		Negative	No studies available		Positive: +	Positive: +++	Positive: +++	Negative	Positive: ++	Negative
Integrated conclusion				derived result						
		Molecular Profiling Related to the Analogue Approach Hypothesis								
		Source 1	Source 2	Target	Source 3	Source 4	Source 5	Outlier 1	Outlier 2	Outlier 3
CAS		88-09-5	97-61-0	4536-23-6	149-57-5	99-66-1	31080-39-4	1185-39-3	1575-72-0	109-52-4
Name		2-ethyl butyric acid	2-methyl pentanoic acid	2-methyl hexanoic acid	2-ethyl hexanoic acid	2-propyl pentanoic acid	2-propyl heptanoic acid	2-dimethyl pentanoic acid	2-propyl pentenoic acid	pentenoic acid
Parent chemical	Profiler 1 (name, version)									
	Expert system 1 (name, version)									
Metabolites*	Profiler 1 (name, version)									
	Expert system 1 (name, version)									
		Physico-Chemical Data								
		Source 1	Source 2	Target	Source 3	Source 4	Source 5	Outlier 1	Outlier 2	Outlier 3
CAS		88-09-5	97-61-0	4536-23-6	149-57-5	99-66-1	31080-39-4	1185-39-3	1575-72-0	109-52-4

Name	2-Ethyl butyric acid	2-Methyl-pentanoic acid	2-Methyl-hexanoic acid	2-Ethylhexanoic acid	Valproic acid	2-Propyl heptanoic acid	2-dimethyl pentanoic acid	4-ene-VPA	pentenoic acid
Molecular weight	116.08	116.08	130.1	144.12	144.12	172.15	130.187	142.2	102.133
Melting point (deg C; predicted EPISUITE)	15.24	15.24	26.62	37.72	37.72	59.15	24.61	36.44	14.76
Boiling point (deg C; predicted EPISUITE)	195.8	195.8	215.45	234.2	234.2	268.98	207.77	232.83	187.75
logPow (experimental EPISUITE)	1.68	1.8	1.8	2.64	2.75	3.2			1.39
logPow (predicted EPISUITE)	1.98	1.98	2.47	2.96	2.96	3.94	2.43	2.82	1.56
Water solubility (mg/L; experimental)	18000 20°C (41)			2000 20°C (16)	2000 at 20°C (37)	275.6 at 25°C (24)			24000 at 25°C
Water solubility (mg/L; human metabolome)									
pKa (experimental)	4.71 (43)			4.7 (19)	4.6 (39)				
pKa (predicted ACD/Percepta)	4.8	4.8	4.8	4.8	4.8	4.8	4.9	4.7	4.84
Vapour pressure (mm Hg; experimental)	0.188 at 25°C (44)			0.03 (18)	0.0458 at 25°C (40)	0.0048 (26)			
Henry's Law Constant (Pa m3/mol; experimental EPISUITE)				0.289					
Henry's Law Constant (Pa m3/mol; predicted EPISUITE (bond method))	0.162	0.172	0.229	0.304	0.304	0.535	0.229	0.226	0.130
	ADME-Toxicokinetics								
	Source 1	Source 2	Target	Source 3	Source 4	Source 5	Outlier 1	Outlier 2	Outlier 3
CAS	88-09-5	97-61-0	4536-23-6	149-57-5	99-66-1	31080-39-4	1185-39-3	1575-72-0	109-52-4
Name	2-Ethyl butyric acid	2-Methyl-pentanoic acid	2-Methyl-hexanoic acid	2-Ethylhexanoic acid	Valproic acid	2-Propyl heptanoic acid	2-dimethyl pentanoic acid	4-ene-VPA	pentenoic acid
Fraction unbound in plasma (human; predicted – Model 1 Mirko-all in domain)	0.348	0.348	0.296	0.138	0.141	0.135	0.282	0.210	0.431

Fraction unbound in plasma (human; predicted – Model 2 Mirko-all in domain)	0.46	0.460	0.382	0.245	0.310	0.191	0.471	0.338	0.434
Fraction unbound in plasma (human; predicted – Model 3 Mirko-all in domain)	0.413	0.359	0.265	0.155	0.193	0.081	0.268	0.307	0.592
Fraction unbound in plasma (human; predicted) - Model 4 (Lhasa-all in domain)	0.712	0.671	0.454	0.201	0.138	0.263	NA	NA	NA
Intrinsic Hepatic Clearance ($CL_{int,H}$; µl/min/10^6 Hepatocytes) (human; experimental Cyprotex)	9.620	10.200	3.950	0.551	0.219	NA	NA	NA	NA
***In vivo* Clearance (CL; L/h) (human; predicted Simcyp, min fu)**	20.42	21.09	9.10	0.85	0.33	NA	NA	NA	NA
***In vivo* Clearance (CL; L/h) (human; predicted Simcyp, max fu)**	29.09	29.07	13.68	1.80	0.73	NA	NA	NA	NA
***In vivo* Clearance (CL; mL/min) (mouse; predicted allometric scaling, min fu)**	0.795	0.821	0.354	0.033	0.013	NA	NA	NA	NA
***In vivo* Clearance (CL; mL/min) (mouse; predicted allometric scaling, max fu)**	1.132	1.131	0.532	0.070	0.028	NA	NA	NA	NA
Steady-state volume of distribution, human (Vss; L/kg; predicted)	0.16	0.16	0.15	0.14	0.14	0.12	0.20	0.13	0.35
Steady-state volume of distribution, human (Vu,ss; L/kg; predicted)	0.35	0.35	0.40	0.57	0.45	0.63	0.33	0.39	0.15
Steady-state volume of distribution, mouse (Vss; L/kg; predicted)	0.30	0.17	0.27	0.13	0.14	0.11	0.40	0.24	0.16
Steady-state volume of distribution, mouse (Vu,ss; L/kg; predicted)	0.65	0.37	0.70	0.53	0.45	0.59	0.63	0.70	0.37

Intrinsic clearance -Insphero										
Supporting data related to the target endpoint(s)										
		Source 1	Source 2	Target	Source 3	Source 4	Source 5	Outlier 1	Outlier 2	Outlier 3
CAS		88-09-5	97-61-0	4536-23-6	149-57-5	99-66-1	31080-39-4	1185-39-3	1575-72-0	591-80-0
Name		EBA	MPA	MHA	EHA	VPA	PHA	DMPA	4-ene-VPA	PA
In vitro	**ZET assay**									
	Effect: (most sensitive) pericardial and/or yolk oedema									
	morphological observation									
	24-48 hpf $EC_{10\ nominal}$ (µM)	363	373	385	65	38	10	426	Not available yet	568
	Effect: small eyes									
	morphological observation									
	24-48 hpf $EC_{10\ nominal}$ (µM)	(1000)[1]	>900[2]	>1000	(400)[1]	(400)[1]	153	540	Not available yet	>1400
	Effect: jitter/tremor									
	behavioural observation									
	96-120 hpf $EC_{10\ nominal}$ (µM)	>1000	>900	442	421	82	16	977	Not available yet	>1400
	Effect: craniofacial deformation									
	morphological observation									
	96-120 hpf $EC_{10\ nominal}$ (µM)	>1000	380	>1000	102	52	21	481	Not available yet	626
	Effect: scoliosis/lordosis									
	morphological observation									
	96-120 hpf $EC_{10\ nominal}$ (µM)	892	393	383	651	59	41	663	Not available yet	747
	ZET Reporter assay									
	CHA angle									

morphological observation									
120 hpf $EC_{10\ nominal}$ (µM)	856	194	N/A	11.9	1.7	0.95	174	6.7	N/A
Conclusion	R^2: 0.54	R^2: 0.75		R^2: 0.77	R^2: 0.75	R^2: 0.91	R^2: 0.58	R^2: 0.94	
mEST assay									
ID_{10} D3 (µM)	>3000	>3000	913	547	275	278	>3000	328	72
ID_{50} D3 (µM)	>3000	>3000	1166	1115	378	365	>3000	518	86
IC_{10} D3 (µM)	2064	2713	>10000	<312.5	415	294	>3000	343	<78
IC_{50} D3 (µM)	>3000	>3000	>10000	4234*	1173	714	>3000	4385	872
IC_{10} 3T3 (µM)	>3000	1838	6834	9568	>3000	1084	>3000	354	79
IC_{50} 3T3 (µM)	>3000	>3000	>10000	10084*	>3000	1873	>3000	>10000	136*
HDAC inhibition EC_{10} (µM) day 4	4640	7380	>10000	340	50	40	>10000	390	950
HDAC inhibition EC_{10} (µM) day 10	2130	1500	2400	160	40	30	3680	50	30
UKN1 assay									
EC_{10} viability (µM)	>5000	>5000	>5000	417	575	263	2042	363	1807
Teratogenicity Score	1,1	6.3	7.9	16.3	17.3	11.9	9.6	17.5	17.6
Classification	0	0	0	1	1	0.5	0.5	1	1
HDAC inhibition EC_{10} (µM)	1040	1190	2370	160	90	30	1240	110	90
CALUX assays EC_{10} (µM)									
Cytotox CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
ERa CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
Anti-ERa CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
AR CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
Anti-AR CALUX	>1000	>1000	>1000	400	>1000	316	316	>1000	>1000
PR CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
Anti-PR CALUX	>1000	>1000	>1000	1000	794	251	>1000	>1000	>1000
GR CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
Anti-GR CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
TRb CALUX	>1000	>1000	>1000	>1000	1000	316	>1000	1000	400
Anti-TRb CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
RAR CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000

	LXR CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
	PXR CALUX	>1000	>1000	>1000	251	79	50	>1000	316	316
	PPARa CALUX	316	316	32	100	100	32	200	79	316
	PPARd CALUX	>1000	>1000	>1000	>1000	>1000	1000	>1000	>1000	>1000
	PPARg CALUX	>1000	>1000	>1000	>1000	>1000	126	>1000	>1000	>1000
	AhR CALUX	>1000	>1000	>1000	>1000	>1000	1000	>1000	>1000	>1000
	Hif1a CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
	TCF CALUX	>1000	>1000	>1000	316	100	50	>1000	316	316
	AP1 CALUX	>1000	>1000	>1000	>1000	>1000	100	>1000	>1000	>1000
	ESRE CALUX	>1000	>1000	>1000	1000	100	100	>1000	1000	1000
	NFkB CALUX	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000	>1000
	Nrf2 CALUX	>1000	>1000	>1000	1000	>1000	100	>1000	>1000	631
	p21 CALUX	>1000	>1000	>1000	1000	631	200	>1000	794	>1000
	p53 GENTOX CALUX	>1000	>1000	>1000	794	251	100	>1000	501	316
1) '(..)' means: response at highest tested and indicated dose <10%; 2) '>'means: not detectable up to the highest tested and indicated dose;										
Supporting data related to the target endpoint(s)										
		Source 1	**Source 2**	**Target**	**Source 3**	**Source 4**	**Source 5**	**Outlier 1**	**Outlier 2**	**Outlier 3**
CAS		88-09-5	97-61-0	4536-23-6	149-57-5	99-66-1	31080-39-4	1185-39-3	1575-72-0	591-80-0
Name		EBA	MPA	MHA	EHA	VPA	PHA	DMPA	4-ene-VPA	PA
In vitro	**ZET assay**									
	Effect: (most sensitive) pericardial and/or yolk oedema									
	morphological observation nominal EC_{10} (µM)	363	373	385	65	38	10	426	Not available	568
	24-48 hpf EC_{10} embryo-total, model (µM)	61	93	96	13	8	1	107	Not available	142
	Effect: small eyes									
	morphological observation nominal EC_{10} (µM)	(1000)[1]	>900[2]	>1000	(400)[1]	(400)[1]	153	540	Not available	>1400
	24-48 hpf EC_{10} embryo-total, model (µM)	(167)[1]	>225	>250	(80)[1]	(80)[1]	19	135	Not available	>350

Effect: jitter/tremor									
behavioural observation nominal EC_{10} (µM)	>1000	>900	442	421	82	16	977	Not available	>1400
96-120 hpf EC_{10}embryo-total, model (µM)	>167	>225	111	84	16	2	244	Not available	>350
Effect: craniofacial deformation									
morphological observation nominal EC_{10} (µM)	>1000	380	>1000	102	52	21	481	Not available	626
96-120 hpf EC_{10}embryo-total, model (µM)	>167	95	>250	20	10	3	120	Not available	157
Effect: scoliosis/lordosis									
morphological observation nominal EC_{10} (µM)	892	393	383	651	59	41	663	Not available	747
96-120 hpf EC_{10}embryo-total, model (µM)	144	98	96	130	12	5	166	Not available	187
1) '(..)' means: response at highest tested and indicated dose <10%; 2) '>'means: not detectable up to the highest tested and indicated dose;									
ZET Reporter assay									
CHA angle									
morphological observation nominal EC_{10} (µM)	856	194	N/A	12	1.7	1	174	7	N/A
120 hpf EC_{10}embryo-total, model (µM)	143	49		2	0.3	0.1	44	1	
mEST assay									
ID_{10} D3 µM, medium-unbound			875	406	189	116		212	71
ID_{50} D3 µM, medium-unbound			1118	827	259	152		336	85
IC_{10} D3 µM, medium-unbound	2000	2601			285	123		222	
IC_{50} D3 µM, medium-unbound				3142	805	298		2841	859

	IC_{10} 3T3 µM, medium-unbound		1762	6552	7101		452		229	78
	IC_{50} 3T3 µM, medium-unbound				7484		781			134
	HDAC inhibition-day 4 EC_{10}, medium-unbound (µM)	4496	7076		252	34	17		253	935
	HDAC inhibition-day 10 EC_{10}, medium-unbound (µM)	2064	1438	2301	119	27	13	3051	32	30
	UKN1 assay									
	EC_{10} viability, medium unbound (µM)				184	215	43	1164	121	1710
	HDAC inhibition EC_{10}, medium-unbound (µM)	931	1028	2048	70	34	5	707	37	85
	CALUX assays EC_{10}cell-total (µM)									
	Anti-AR CALUX				28		32	38		
	Anti-PR CALUX				70	87	25			
	TRb CALUX					110	32		110	48
	PXR CALUX				18	9	5		35	38
	PPARa CALUX	41	41	4	7	11	3	24	9	38
	PPARd CALUX						100			
	PPARg CALUX						13			
	AhR CALUX						100			
	TCF CALUX				22	11	5		35	38
	AP1 CALUX						10			
	ESRE CALUX				70	11	10		110	38
	Nrf2 CALUX				70		10			76
	p21 CALUX				70	69	20		87	
	p53 GENTOX CALUX				56	28	10		55	38
Supporting Data Related to the Target Endpoint(s)										
		Source 1	**Source 2**	**Target**	**Source 3**	**Source 4**	**Source 5**	**Outlier 1**	**Outlier 2**	**Outlier 3**
CAS		88-09-5	97-61-0	4536-23-6	149-57-5	99-66-1	31080-39-4	1185-39-3	1575-72-0	591-80-0
Name		EBA	MPA	MHA	EHA	VPA	PHA	DMPA	4-ene-VPA	PA

ZET Assay									
pericardial and/or yolk oedema EC_{10} Nominal (µM)	363	373	385	65	38	10	426	N/A	568
human oral dose equivalent (min fu model) (mg/kg)	66.7	71.4	54.5	8.3	4.9				
human oral dose equivalent (max fu model) (mg/kg)	63	67.7	47.9	5.3	3.1				
mouse oral dose equivalent (min fu model) (mg/kg)	59.2	63	56.9	11	6.7				
mouse oral dose equivalent (max fu model) (mg/kg)	58.7	62.5	47.2	7.3	4.3				
pericardial and/or yolk oedema EC_{10} embryo-total, model (µM)	61	93	96	13	8	1	107	N/A	142
human oral dose equivalent (min fu model) (mg/kg)	51.3	62.3	57.5	6.1	3.6				
human oral dose equivalent (max fu model) (mg/kg)	53.3	84.5	55.1	4.8	2.8				
mouse oral dose equivalent (min fu model) (mg/kg)	8.93	14.1	9.77	0.9	0.55				
mouse oral dose equivalent (max fu model) (mg/kg)	15.6	23.6	13.2	1.2	0.74				
small eyes EC_{10} Nominal (µM)	(1000)[1]	>900[2]	>1000	(400)[1]	(400)[1]	153	540	N/A	>1400
human oral dose equivalent (min fu model) (mg/kg)	183.7			50.9	51.9				
human oral dose equivalent (max fu model) (mg/kg)	173.4			32.6	32.7				
mouse oral dose equivalent (min fu model) (mg/kg)	163			67.5	70.4				
mouse oral dose equivalent (max fu model) (mg/kg)	161.7			44.7	45.3				
small eyes EC_{10} embryo-total, model (µM)	(167)[1]	>225	>250	(80)[1]	(80)[1]	19	135	N/A	>350
human oral dose equivalent (min fu model) (mg/kg)	140.5			37.5	36				
human oral dose equivalent (max fu model) (mg/kg)	146			29.5	27.9				

mouse oral dose equivalent (min fu model) (mg/kg)	24.4			5.45	5.5				
mouse oral dose equivalent (max fu model) (mg/kg)	42.7			7.25	7.4				
jitter/tremor EC_{10} Nominal (µM)	>1000	>900	442	421	82	16	977	N/A	>1400
human oral dose equivalent (min fu model) (mg/kg)			62.6	53.4	10.6				
human oral dose equivalent (max fu model) (mg/kg)			55	34.3	6.7				
mouse oral dose equivalent (min fu model) (mg/kg)			65.4	71	14.5				
mouse oral dose equivalent (max fu model) (mg/kg)			54.2	47	9.3				
jitter/tremor embryo-total, model (µM)	>167	>225	111	84	16	2	244	N/A	>350
human oral dose equivalent (min fu model) (mg/kg)			66.4	39.4	7.2				
human oral dose equivalent (max fu model) (mg/kg)			63.7	30.9	5.6				
mouse oral dose equivalent (min fu model) (mg/kg)			11.7	5.72	1.1				
mouse oral dose equivalent (max fu model) (mg/kg)			15.26	7.61	1.48				
craniofacial deformation EC_{10} Nominal (µM)	>1000	380	>1000	102	52	21	481	N/A	626
human oral dose equivalent (min fu model) (mg/kg)		72.7		13	6.7				
human oral dose equivalent (max fu model) (mg/kg)		69		8.3	4.3				
mouse oral dose equivalent (min fu model) (mg/kg)		64.2		17.2	9.2				
mouse oral dose equivalent (max fu model) (mg/kg)		63.7		11.4	5.9				
craniofacial deformation embryo-total, model (µM)	>167	95	>250	20	10	3	120	N/A	157
human oral dose equivalent (min fu model) (mg/kg)		83		9.4	4.5				

human oral dose equivalent (max fu model) (mg/kg)		86.3		7.36	3.5				
mouse oral dose equivalent (min fu model) (mg/kg)		14.4		1.36	0.69				
mouse oral dose equivalent (max fu model) (mg/kg)		24.1		1.8	0.925				
scoliosis/lordosis EC_{10} Nominal (µM)	892	393	383	651	59	41	663	N/A	747
human oral dose equivalent (min fu model) (mg/kg)	164	75.2	54.2	82.6	7.7				
human oral dose equivalent (max fu model) (mg/kg)	154.6	71.4	47.7	53	4.8				
mouse oral dose equivalent (min fu model) (mg/kg)	145.5	66.4	56.7	109.9	10.4				
mouse oral dose equivalent (max fu model) (mg/kg)	144.2	65.9	46.96	72.7	6.7				
scoliosis/lordosis embryo-total, model (µM)	144	98	96	130	12	5	166	N/A	187
human oral dose equivalent (min fu model) (mg/kg)	121.2	85.7	57.5	61.1	5.4				
human oral dose equivalent (max fu model) (mg/kg)	125.8	89	55.1	47.84	4.2				
mouse oral dose equivalent (min fu model) (mg/kg)	21.1	14.9	10.1	8.86	0.83				
mouse oral dose equivalent (max fu model) (mg/kg)	36.8	24.9	13.2	11.8	1.11				
ZET Reporter assay									
EC_{10} Nominal Treatment Conc. (µM)	856	194	N/A	12	1.7	1	174	7	N/A
human oral dose equivalent (min fu model) (mg/kg)	157.4	37.1		1.5	0.22				
human oral dose equivalent (max fu model) (mg/kg)	148.4	35.2		0.98	0.138				
mouse oral dose equivalent (min fu model) (mg/kg)	139.6	32.8		2.02	0.3				
mouse oral dose equivalent (max fu model) (mg/kg)	138.4	32.5		1.34	0.19				
EC_{10}embryo-total, model (µM)	143	49	N/A	2	0.3	0.1	44	1	N/A

human oral dose equivalent (min fu model) (mg/kg)	120.30	42.70		0.94	0.135				
human oral dose equivalent (max fu model) (mg/kg)	124.90	44.50		0.74	0.105				
mouse oral dose equivalent (min fu model) (mg/kg)	20.90	7.45		0.136	0.020				
mouse oral dose equivalent (max fu model) (mg/kg)	36.50	12.45		0.18	0.028				
mEST assay									
ID_{10} D3 µM (medium-unbound)			875	406	189	116		212	71
human oral dose equivalent (min fu model) (mg/kg)			124.00	51.6	24.5				
human oral dose equivalent (max fu model) (mg/kg)			108.90	33	16				
mouse oral dose equivalent (min fu model) (mg/kg)			129.50	69	33				
mouse oral dose equivalent (max fu model) (mg/kg)			107.20	45.3	21				
IC_{10} D3 µM (medium-unbound)	2000	2601			285	123		222	
human oral dose equivalent (min fu model) (mg/kg)	367.8	497.50			36.9				
human oral dose equivalent (max fu model) (mg/kg)	347.00	472.30			23				
mouse oral dose equivalent (min fu model) (mg/kg)	326.20	439.40			50				
mouse oral dose equivalent (max fu model) (mg/kg)	323.40	436.00			32				
IC_{10} 3T3 µM (medium-unbound)		1762	6552	7101		452		229	78
human oral dose equivalent (min fu model) (mg/kg)		337.0	928.40	901.3					
human oral dose equivalent (max fu model) (mg/kg)		320.00	815.30	578					
mouse oral dose equivalent (min fu model) (mg/kg)		297.70	969.10	1199					
mouse oral dose equivalent (max fu model) (mg/kg)		295.40	802.60	793					

HDAC inhibition-day 4 EC_{10}medium-unbound (µM)	4496	7076		252	34	17		253	935
human oral dose equivalent (min fu model) (mg/kg)	826.6	1353.30		32	4.4				
human oral dose equivalent (max fu model) (mg/kg)	779.20	1284.80		21	3				
mouse oral dose equivalent (min fu model) (mg/kg)	733.30	1195.00		43	6				
mouse oral dose equivalent (max fu model) (mg/kg)	727.00	1186.30		28.1	4				
HDAC inhibition-day 10 EC_{10}medium-unbound (µM)	2064	1438	2301	119	27	13	3051	32	30
human oral dose equivalent (min fu model) (mg/kg)	379.6	275.00	326.10	15	3.49				
human oral dose equivalent (max fu model) (mg/kg)	357.80	261.10	286.30	10	2				
mouse oral dose equivalent (min fu model) (mg/kg)	336.60	242.90	340.30	20	5				
mouse oral dose equivalent (max fu model) (mg/kg)	333.70	241.10	281.90	13.3	3				
UKN1 assay									
Viability EC_{10}medium-unbound (µM)				184	215	43	1164	121	1710
human oral dose equivalent (min fu model) (mg/kg)				23	28				
human oral dose equivalent (max fu model) (mg/kg)				15	18				
mouse oral dose equivalent (min fu model) (mg/kg)				31	38				
mouse oral dose equivalent (max fu model) (mg/kg)				21	24				
HDAC inhibition EC_{10} medium-unbound (µM)	931	1028	2048	70	34	5	707	37	85
human oral dose equivalent (min fu model) (mg/kg)	171.20	196.60	290.20	8.9	4				
human oral dose equivalent (max fu model) (mg/kg)	161.40	186.60	254.80	6	3				

mouse oral dose equivalent (min fu model) (mg/kg)		151.80	173.60	302.90	12	6				
mouse oral dose equivalent (max fu model) (mg/kg)		150.50	172.40	250.90	7.8	4				
CALUX assays										
EC_{10}cell-total (µM)	min.	41	41	4	7	9	3	24	9	38
	max.	41	41	4	70	110	100	38	110	76
human oral dose equivalent (min fu model) (mg/kg)	min.	34.4	35.8	2.4	3.29	3.95				
	max.	34.4	35.8	2.4	32.9	49.5				
human oral dose equivalent (max fu model) (mg/kg)	min.	35.8	37.3	2.3	2.6	3.2				
	max.	35.8	37.3	2.3	25.7	38.3				
mouse oral dose equivalent (min fu model) (mg/kg)	min.	6	6.2	0.42	0.48	0.62				
	max.	6	6.2	0.42	4.8	7.6				
mouse oral dose equivalent (max fu model) (mg/kg)	min.	10.5	10.4	0.55	0.63	0.83				
	max.	10.5	10.4	0.55	6.4	10.14				
In chemico										
		Source 1	**Source 2**	**Target**	**Source 3**	**Source 4**	**Source 5**	**Outlier 1**	**Outlier 2**	**Outlier 3**
CAS		88-09-5	97-61-0	4536-23-6	149-57-5	99-66-1	31080-39-4	1185-39-3	1575-72-0	591-80-0
Name		EBA	MPA	MHA	EHA	VPA	PHA	DMPA	4-ene-VPA	PA
		In silico models of OECD QSAR Toolbox								
DNA binding by OASIS		No alert found	No alert found	No alert found	No alert found	No alert found	No alert found			
DNA binding by OECD		No alert found	No alert found	No alert found	No alert found	No alert found	No alert found			
Eye irritation/corrosion Exclusion rules by BfR		Undefined	Undefined	Undefined	Undefined	Undefined	Group C Melting Point > 55 C|Undefined			
Eye irritation/corrosion Inclusion rules by BfR		Inclusion rules not met	Inclusion rules not met	Inclusion rules not met	Inclusion rules not met	Inclusion rules not met	Inclusion rules not met			
Protein Binding Potency h-CLAT		No alert found	No alert found	No alert found	No alert found	No alert found	No alert found			

Protein binding alerts for Chromosomal aberration by OASIS	No alert found	No alert found	No alert found	No alert found	No alert found	No alert found			
Protein binding alerts for skin sensitization according to GHS	No alert found	No alert found	No alert found	No alert found	No alert found	No alert found			
Protein binding alerts for skin sensitization by OASIS	No alert found	No alert found	No alert found	No alert found	No alert found	No alert found			
Protein binding by OASIS	No alert found	No alert found	No alert found	No alert found	No alert found	No alert found			
Protein binding by OECD	No alert found	No alert found	No alert found	No alert found	No alert found	No alert found			
Protein binding potency GSH	Not possible to classify	Not possible to classify	Not possible to classify	Not possible to classify	Not possible to classify	Not possible to classify			
Protein binding potency Cys (DPRA 13%)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)			
Protein binding potency Lys (DPRA 13%)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)	Non-Conjugated carboxylic acids and esters (non reactive)			
Skin irritation/corrosion Exclusion rules by BfR	Undefined	Undefined	Undefined	Undefined	Undefined	Group C Melting Point > 55 C|Undefined			
Skin irritation/corrosion Inclusion rules by BfR	Aliphatic acids	Aliphatic acids	Aliphatic acids	Aliphatic acids	Aliphatic acids	Inclusion rules not met			

Annex II. *In vitro* models - detailed description of methods and generated data

Please refer to the separate publication for full Annex II

ENV/JM/MONO(2020)21/ANN2

Annex III. *In silico* models - detailed description of methods and generated data

Please refer to the separate publication for full Annex III

ENV/JM/MONO(2020)21/ANN3