



Modélisation pharmacocinétique et pharmacodynamique de la toxicité développementale de l'embryon de poisson zèbre.

Ségolène Siméon

► To cite this version:

Ségolène Siméon. Modélisation pharmacocinétique et pharmacodynamique de la toxicité développementale de l'embryon de poisson zèbre.. Biologie de la reproduction. Institut agronomique, vétérinaire et forestier de France, 2020. Français. NNT : 2020IAVF0020 . tel-03185019

HAL Id: tel-03185019

<https://pastel.hal.science/tel-03185019>

Submitted on 30 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT

préparée à l'Institut des sciences et industries du vivant et de l'environnement (AgroParisTech)

pour obtenir le grade de

Docteur de l'Institut agronomique vétérinaire et forestier de France

Spécialité : Toxicologie

École doctorale n° 581

Agriculture, alimentation, biologie, environnement et santé (ABIES)

par

Ségolène SIMEON

Modélisation pharmacocinétique et pharmacodynamique de la toxicité développementale de l'embryon de poisson zèbre

Directeur de thèse : Frédéric BOIS

Co-encadrement de la thèse : Rémy BEAUDOUIN

Thèse présentée et soutenue à Paris, le 9 décembre 2020 :

Composition du jury :

M. Eric PARENT, IGPEF, AgroParisTech

Mme Marie-Laure DELIGNETTE-MULLER, Professeur, VetAgro Sup

M. Frédéric ALONZO, Chercheur (HDR), IRSN

M. Jean-Lou DORNE, Directeur de recherche, EFSA

M. Iñaki ITURRIA, R&D Project manager, Biobide

M. Alexandre PERY, ICPEF, AgroParisTech

M. Rémy BEAUDOUIN, Ingénieur de recherche (HDR), Ineris

M. Frédéric BOIS, Chercheur (HDR), Certara

Président

Rapporteur & Examinatrice

Rapporteur & Examineur

Examineur

Examineur

Examineur

Co-encadrant & Examineur

Directeur de thèse

REMERCIEMENTS

Mes premiers remerciements, et non plus des moindres, sont évidemment pour mon directeur de thèse Frédéric Bois. Merci d'avoir su m'accompagner pendant ces trois années quelque peu secouées par divers challenges mais surtout enrichies de rencontres, congrès, découvertes, apprentissages et expériences que tu m'as fait vivre. Tu m'as fait retrouver le goût pour la modélisation au moment où le doute a essayé de me faire flancher. Malgré ton départ de l'INERIS, tu ne m'as pas abandonnée, tu es resté incroyablement disponible pour moi et tu as su trouver le moyen de m'encadrer à la perfection, avec humour et franchise. Tu m'as écoutée, (re)motivée, raisonnée, bottée les fesses quand il le fallait (tête en l'air que je suis...), félicitée aussi quand je le méritais, dans le but de me faire avancer et accomplir ce beau projet que j'aurais à coup sûr regretté d'avoir abandonné. Je ne trouve pas de mots plus fort que *Merci beaucoup*, alors je te dédie avec fierté mes travaux qui sans toi n'auraient pu être menés à bien. J'espère que nos routes continueront de se croiser, ce qui est fort probable grâce à tes précieux conseils et appuis pour me construire un bel avenir.

Je remercie grandement toute l'équipe METO au sein de laquelle j'ai évolué. Merci à Céline Brochot d'avoir mis tout en œuvre pour que ma thèse se déroule dans les meilleures conditions. Merci à Cleo Tebby, Florence Zeman et Enrico Mombelli pour les joyeuses discussions partagées. Merci également à Vanessa Veg et Caroline Gourland pour leur disponibilité et sympathie. Pour finir, un spécial grand merci à Rémy Beaudouin pour avoir suivi et encadré mes travaux de près, ainsi que pour le soutien apporté.

Je souhaite également remercier tous les partenaires EU-ToxRisk avec qui j'ai apprécié de travailler, Ciaran Fisher, Ian Gardner, Dinant Kroese et Katharina Brotzmann.

Je remercie les membres de mon comité de thèse, François Brion, Frédérique Clément, et Celia Quevedo pour leurs avis et recommandations sur mes travaux de thèse. Un grand merci à Sandrine Micallef qui, au-delà de ce rôle, m'a guidée pour statuer sur mes envies personnelles et professionnelles pour m'orienter vers l'avenir scientifique qui m'est destiné.

Un immense merci à mes amis doctorants, Paul, Audrey, Marc, Viviane et Alice, mes compagnons de galère et d'aventures. Sans vous, ces trois années n'auraient pas eu la même consonnance. Du choix du morceau de pain le plus biscornu à la cantine, en passant par des parties de tennis endiablées, des musiques agaçantes qui restent dans la tête des jours et des

jours, ou encore des discussions profondément profondes, le train TER à destination du PhD, initialement prévu pour retrouver ses collègues, a été dévié vers l'amitié par vos merveilleuses personnalités, parfois étranges...

Un infini de mercis à l'incroyable personnage que la pharmacocinétique m'a permis de rencontrer, Alix. Un mois à peine a suffi pour comprendre qu'on ne se quitterait plus. Tu es bien plus qu'une amie pour moi. Tu es LA personne qui a su trouver les mots quand il le fallait, avec bienveillance et sincérité. Tu es mon plus grand soutien, l'Ange gardien qui m'a remis sur le droit chemin. Tu es aussi mon modèle, te voir accomplir avec humilité et modestie tes brillants projets me poussent à ne pas seulement me contenter du minimum mais à viser l'excellence. Après tout, il y a de la place pour deux sur le tapis rouge. Je te remercie aussi énormément de m'avoir un jour forcé à partir en vacances avec tes plus aimables amis Margaux et Floflo. Pendant que tu brilles en société à Berlin, je me suis permise de les accaparer un petit peu car leur présence m'est devenue indispensable pour garder un moral sans faille.

Un grand merci pour ceux avec qui cette carrière de scientifique a commencé, mes amis de P1, Emilie, Mégane, Aliette, Alex et Gwen. Puis ceux avec qui elle s'est concrétisée, mes quatre doigts de la main, Charlotte, Margue, Clem et Margaux, ainsi que tous les autres qui ont rendus mes années de pharma inoubliables, Maxime, Agathe, Amélie, Clémentine, Amélie, Maëlle, Julie, Paul... Je remercie aussi respectueusement les non scientifiques qui ont subi à leurs dépens mes apitoiements pendant ces années, Aleks, Solange et mon Carolo. A huge thank you à Thea pour ces amazing bons moments passés ensemble à refaire le monde, orchestrés par une imagination farfelue débordante. Même à l'autre bout de la planète, ton soutien et tes encouragements quotidiens si réconfortants m'ont été d'une grande aide. Muchas gracias et un enorme grazie à mes fidèles amis Andrea et Giancarlo. Ne changez jamais, la vie avec vous est si joyeuse et belle, les moments partagés à vos côtés sont tout simplement parfaits.

Pour finir, j'adresse mes sincères remerciements à toute ma famille pour les moments partagés ensemble, si important pour s'octroyer une bonne bouffée d'air frais et trouver le courage de repartir.

TABLE DES MATIERES

LISTE DES FIGURES ET DES TABLEAUX.....	6
ABREVIATIONS.....	8
INTRODUCTION GÉNÉRALE	11
CHAPITRE I. Etat de l'art	14
1. <i>La modélisation mathématique pour l'extrapolation à l'Homme en évaluation des risques</i>	14
1.1. La modélisation pharmacocinétique.....	15
1.1.1. Les modèles compartimentaux classiques	16
1.1.2. Les modèles physiologiques	17
1.2. La modélisation pharmacodynamique.....	19
1.2.1. Les modèles empiriques	19
1.2.2. Les modèles mécanistes	21
1.2.3. Les modèles intermédiaires entre empiriques et mécanistes	24
1.3. L'inférence statistique sur les paramètres des modèles	29
1.3.1. L'inférence bayésienne.....	29
1.3.2. Les modèles multi-niveaux	30
1.3.3. La calibration des modèles développés pour les travaux de thèse	31
1.4. Les intérêts et limites de la modélisation pour l'extrapolation à l'Homme	32
2. <i>Les données de toxicologie développementale et reproductive</i>	34
2.1. Les données <i>in vivo</i>	34
2.2. Les limites des données <i>in vivo</i> pour l'extrapolation à l'Homme	35
2.3. Les données <i>in vitro</i>	35
2.4. Les limites des données <i>in vitro</i> pour l'extrapolation à l'Homme	36
2.5. Les tests sur embryons de poisson zèbre (ZET)	37
2.5.1. Caractéristiques de l'espèce.....	37
2.5.2. Intérêt pour étudier la toxicité développementale.....	38
2.5.3. Protocole d'étude de l'OCDE	41

2.5.4. Les limites des données des tests sur les embryons de poisson zèbre ...	42
2.5.5. Modèles mathématiques pour le poisson zèbre	43
2.5.6. Synthèse des ZET	45
3. <i>Les composés cibles en toxicologie développementale</i>	45
CHAPITRE II. Modèle de distribution <i>in vitro</i> VIVD.....	50
1. <i>Un nouveau modèle de distribution virtuelle in vitro</i>	50
1.1. Objectif des travaux et contribution	50
1.2. Perspectives.....	51
2. <i>Résumé et article du modèle VIVD</i>	51
CHAPITRE III. Modèle PBPK pour évaluer de la toxicité développementale chez l'embryon de poisson zèbre.....	76
1. <i>Un nouveau modèle PBPK générique pour l'embryon de poisson zèbre</i>	76
1.1. Objectif des travaux et contribution	76
1.2. Articulation avec le modèle VIVD précédent	77
1.3. Nécessité d'un corrigé.....	78
1.4. Perspectives.....	78
2. <i>Résumé et article du modèle PBPK</i>	79
CHAPITRE IV. Modèle PK-PD multi-états pour étudier les effets des composés chimiques .	119
1. <i>Un nouveau modèle multi-états de la toxicité développementales des composés chimiques chez l'embryon de poisson zèbre</i>	119
1.1. Objectif des travaux et contribution	119
1.2. Articulation avec le modèle PBPK précédent.....	120
1.3. Perspectives.....	120
2. <i>Résumé et article du modèle PK-PD multi-états</i>	121
DISCUSSION GENERALE ET PERSPECTIVES	154
1. <i>L'importance des données expérimentales pour un modèle prédictif pertinent</i>	155
1.1. L'élaboration du test expérimental.....	156
1.2. L'analyse des données et ajustement du modèle.....	157
1.3. La validation du modèle	158
2. <i>Les limites et perspectives d'amélioration des modèles pour prédire la toxicité développementale de l'embryon de poisson zèbre</i>	158

2.1. Le modèle PBPK.....	159
2.1.1. Les coefficients de partage.....	159
2.1.2. L'hypothèse de diffusion instantanée dans l'embryon.....	160
2.1.3. Les volumes tissulaires.....	160
2.1.4. La clairance métabolique.....	161
2.1.5. Les concentrations effectives.....	161
2.1.6. Comparaison avec les modèles de la littérature.....	162
2.2. Le modèle multi-états.....	164
2.2.1. La structure du modèle.....	164
2.2.2. La calibration du modèle.....	166
2.2.3. Les concentrations effectives.....	167
3. <i>La place des modèles mécanistes dans l'évaluation des risques</i>	168
CONCLUSION GÉNÉRALE.....	171
REFERENCES.....	172
ANNEXES.....	186
<i>Annexe 1 : Supplementary Material (Article modèle VIVD).....</i>	<i>186</i>
<i>Annexe 2 : Supporting Information (Article modèle PBPK).....</i>	<i>187</i>
<i>Annexe 3 : Supplementary Material (Article modèle PK-PD multi-états).....</i>	<i>214</i>

LISTE DES FIGURES ET DES TABLEAUX

Les figures

- Figure 1. Schéma d'un modèle pharmacocinétique mono (A) et bi-compartimental (B), avec Q_1 et Q_2 les quantités du composé dans les compartiments, k_a et k_e respectivement les constantes de vitesse d'absorption et d'élimination, $k_{1,2}$ et $k_{2,1}$ les constantes de vitesse de transfert. 16
- Figure 2. Modèle pharmacocinétique basé sur la physiologie du poisson adulte. L'absorption, l'élimination et le métabolisme sont représentés par les flèches orange, bleue et verte, respectivement (Grech et al., 2019)..... 18
- Figure 3. Schéma d'un modèle PBPK/PD de l'acide perfluorooctanesulfonique (Sharma et al., 2017)..... 22
- Figure 4. Schéma du modèle de la voie du métabolisme du GSH par Reed et al. (2008) 23
- Figure 5. Représentation schématique des différents niveaux organisationnels d'un AOP, adapté à partir de (Ankley et al., 2010). 24
- Figure 6. Organisation et utilisation des chemins de l'effet néfaste (AOP) pour l'évaluation des risques (Perkins et al., 2019). 25
- Figure 7. Schéma du processus basique d'analyse de survie. 26
- Figure 8. Schéma d'un processus simple des risques concurrents. 27
- Figure 9. Schéma d'un modèle multi-états, avec événements intermédiaires et possibilité de retour entre états. 28
- Figure 10. Principales données nécessaires à l'extrapolation quantitative des données à l'Homme et de la prédiction de la toxicité. Les prédictions (en bleu) de la pharmacocinétique *in vivo*, des effets cellulaires *in vivo*, ainsi que des paramètres pathologiques (toxicité), sont obtenues par différents modèles (en vert) capables d'intégrer des données (en rouge) provenant de diverses sources (Nadia Quignot et al., 2014). 33
- Figure 11. Cycle de vie du poisson zèbre (D'Costa and Shepherd, 2009)..... 38
- Figure 12. Stades de vie des poissons, statut de protection et durée des tests de toxicité standard. Les flèches pleines désignent la durée des tests référencés; la flèche en pointillés indique la durée supplémentaire autorisée pour les tests de l'embryon de poisson «étendu», comprenant le stade éléuthéro-embryonnaire. Adapté de (Embry et al., 2010)..... 42

Figure 13. Structure chimique de l'acide valproïque.	47
Figure 14. Mécanisme d'action pharmacologique de l'acide valproïque (Chuang et al., 2012).	48
Figure 15. Le carrousel de la modélisation, d'après Gabrielsson et Weiner (2006).	156
Figure 16. Courbes dose-réponse sur la base des concentrations nominales pour le VPA et ses analogues (d'après les travaux de K. Brotzmann).....	161
Figure 17. Stades précoces de développement d'inclusion de l'embryon de poisson zèbre pour les tests de toxicité, de 0,2 –1,75 heures post-fertilisation (Kimmel et al., 1995b).	165

Les tableaux

Tableau 1. Tableau résumé des principaux domaines d'application des modèles PK, PBPK et PK/PD développés pour le poisson zèbre.....	44
--	----

ABREVIATIONS

ADME : Absorption, distribution, métabolisme et excrétion

AIC : Critère d'information d'Akaike

Anses : Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail

AO : Effet néfaste (Adverse Outcome)

(q)AOP : Chemin de l'effet néfaste (quantitatif) (Adverse Outcome Pathway)

BCF : Facteur cinétique de biotransformation

CL : Concentration létale

DART : tests de toxicité pour le développement et la reproduction

DL : Dose létale

EC : Concentration effective

EDO : Equation différentielle ordinaire

FET : Test sur l'embryon de poisson (Fish Embryo Test)

GABA : Acide γ -aminobutyrique

IATA : Approches intégrées pour tester et évaluer

KE : Evènement clé (Key Event)

LOAEL : Dose minimale pour observer un effet toxique (Lowest observed adverse effect level)

MIE : Evènement moléculaire initial (Molecular Initial Event)

NAM : Nouvelles méthodes d'approche

NOAEL : Dose sans effet toxique observable (No observable adverse effect level)

NTD : Défaut du tube neural

OCDE : Organisation de coopération et de développement économiques

OMS : Organisation mondiale de la santé

PBPK : Pharmacocinétique basée sur la physiologie

PFOS : Acide perfluorooctanesulfonique

PD : Pharmacodynamique

PK : Pharmacocinétique

QSAR : Relation quantitative structure – activité

REACH : Enregistrement, évaluation et autorisation des substances chimiques (Registration, Evaluation and Authorisation of Chemicals)

TCA : Cycle de l'acide tricarboxylique

TK – TD : Toxicocinétique – toxicodynamie

VPA : Acide valproïque

WB-PBPK : Pharmacocinétique basée sur la physiologie – corps entier

ZET : Test sur les embryons de poisson zèbre

INTRODUCTION GÉNÉRALE

Comprendre et prédire les effets des substances chimiques sur le développement des êtres vivants représente en toxicologie un réel défi afin d'appréhender les problématiques environnementales et thérapeutiques. L'évaluation des risques est une discipline scientifique de santé dont une des priorités est de caractériser la toxicité des composés pour extrapoler quantitativement les effets à l'Homme avec un fort degré de certitude. Elle est constituée de quatre étapes : l'identification des dangers, l'évaluation de la relation dose – réponse et de l'exposition, ainsi que la caractérisation des risques (Becking, 1995). Dans un premier temps, l'identification des dangers consiste à déterminer la potentialité d'un composé à provoquer un effet indésirable. Ensuite, l'évaluation de la dose – réponse et de l'exposition établit le mécanisme d'action ainsi que la quantité de composé nécessaire pour provoquer un effet indésirable. Enfin, la caractérisation des risques intègre les informations sur le danger, la relation dose – réponse, les expositions (externes) du corps entier et les expositions (internes) des cellules cibles de la toxicité pour estimer la probabilité d'effets nocifs chez les individus et/ou les populations exposées (Perkins et al., 2019).

Depuis quelques années, les progrès scientifiques et techniques, ainsi que l'évolution des considérations éthiques concernant l'expérimentation animale remettent en cause les méthodes traditionnelles de l'évaluation des dangers, des expositions internes et des relations dose – réponses, en prônant une vision holistique des sciences biologiques et toxicologiques. Les méthodes les plus courantes passent par des expérimentations animales coûteuses, soulevant des problèmes de fiabilité, de reproductibilité et de compréhension des mécanismes d'action des composés. Pour les remplacer, des méthodes de calcul mathématique (*in silico*), et de biologie moléculaire ou cellulaire (*in vitro*) sont activement développées. L'analyse des tests menés *in vivo* est également améliorée. L'objectif est d'accroître la compréhension des déterminants des expositions internes (pharmacocinétique ou toxicocinétique) et des mécanismes de toxicité (pharmacodynamie ou toxicodynamie). Le critère mécaniste est primordial pour prétendre à des extrapolations quantitatives pertinentes des données de toxicité *in vitro* ou animales à l'Homme (Margiotta-Casaluci et al., 2016; Punt, 2020). Ultimement, ces méthodes devraient fournir des prédictions précises des effets chez l'humain, afin d'établir des liens de causalité et prévenir les effets néfastes suite à l'exposition aux composés chimiques,

tout en satisfaisant le principe des 3R (remplacement, réduction, raffinement) (Beilmann, 2018; Lanzoni et al., 2019).

Parmi les méthodes alternatives disponibles, les tests sur l'embryon de poisson zèbre (*Danio rerio*, assimilé par la Commission européenne à un système *in vitro*) sont de plus en plus utilisés pour étudier divers types de toxicité. L'espèce présente des homologues génétiques avec l'Homme, n'est pas contrainte aux réglementations concernant le bien-être animal lors des premières heures de vie et offre de nombreux avantages expérimentaux, à savoir, une facilité d'élevage due à une large tolérance aux conditions environnementales, une fécondité élevée, une petite taille, un développement rapide ainsi que la transparence des embryons et larves rendant possibles les explorations non invasives (Braunbeck et al., 2015; Cassar et al., 2020; OECD, 2013).

Nous avons identifié deux limites importantes à l'analyse et l'utilisation des tests sur embryon de poisson zèbre pour prédire la toxicité développementale humaine :

- Premièrement, il n'existait pas, avant nos travaux, de modèle mathématique ou statistique capable d'intégrer et d'expliquer les données sur l'exposition interne des cellules de cet embryon, ou de simuler cette exposition pour des scénarios d'exposition externe arbitraires ;
- Deuxièmement, toutes les analyses statistiques des données d'occurrence de malformations et mortalité chez cet embryon faisaient, avant nos travaux, abstraction du facteur temps et de la non-indépendance des données sur les différents effets.

Pour pallier ces manques, nous avons donc :

- Développé un modèle pharmacocinétique physiologique (modèle PBPK) capable de prédire les concentrations de composés neutres ou ionisables dans les cellules et organelles cellulaires (mitochondries et lysosomes) de l'embryon de poisson zèbre.
- Développé un modèle pharmacocinétique – pharmacodynamique (PK-PD) multi-états capable de prédire la probabilité d'occurrence de diverses malformations chez l'embryon, en fonction du temps et des concentrations d'expositions internes en acide valproïque (VPA).

Initié en 2016, le projet européen *EU-ToxRisk* a pour but de développer et tester de nouvelles méthodes alternatives à l'expérimentation animale pour s'assurer de la sécurité des produits chimiques. Les travaux de cette thèse font partie intégrante d'une étude de cas et de deux tâches de développement méthodologique pour la prédiction de la toxicité

développementale humaine du VPA (tératogène notoire) et neuf de ses analogues, sur la base d'essais chez l'embryon de poisson zèbre.

Le premier chapitre présente l'état de l'art sur les questions abordées. Le second est consacré à la description du modèle mathématique de partage des molécules entre le milieu de culture et les cellules utilisé pour l'obtention des paramètres d'entrée du modèle PBPK. L'élaboration et les résultats des deux modèles PBPK et PK-PD multi-états sont détaillés dans les chapitres trois et quatre. Finalement, les limites, les retombées et les enjeux futurs de ce travail sont discutés en dernière partie.

CHAPITRE I. Etat de l'art

Ce chapitre fait un état des connaissances sur les thèmes abordés au cours de ce travail de recherche. La première partie présente les différents modèles pharmacocinétiques et pharmacodynamiques utilisés comme outils pour l'extrapolation à l'Homme en évaluation des risques. Essentielles à l'élaboration de modèles robustes, les grandes lignes des protocoles d'obtention de données de toxicologie développementale sont exposées en deuxième partie. Pour finir, la troisième partie décrit succinctement les principaux composés dont le potentiel toxique sur le développement est notoire ou suspecté.

1. La modélisation mathématique pour l'extrapolation à l'Homme en évaluation des risques

Depuis quelques années, les investigations toxicologiques se sont tournées vers une approche fondée sur des preuves, plutôt que descriptive, notamment pour contrer les coûts élevés, la durée des études *in vivo*, ainsi que les strictes réglementations concernant le bien-être animal (Beilmann, 2018). Pour des raisons éthiques, les tests de toxicité ne peuvent être réalisés directement chez les humains. Différents outils *in vitro*, *in vivo* et *in silico* ainsi que des protocoles d'expérimentations animales ont donc dû être développés afin d'étudier indirectement la toxicité humaine des composés chimiques qui nous entourent. Cette partie a pour objectif de présenter les modèles pharmacocinétiques et pharmacodynamiques utilisés en évaluation des risques pour étudier la toxicité humaine.

L'évaluation des composés repose sur des modèles de devenir, d'exposition, de toxicité, de risque et d'impact. Pour la gestion des risques, le développement conjoint de ces modèles, visant à améliorer l'interprétation et l'analyse des données, permet d'apporter un soutien scientifique pour la prise de décisions réglementaires (McCarty et al., 2018). Un modèle mathématique pertinent est une simplification conceptuellement plausible d'un processus ou d'une activité basée sur des données expérimentales. Ces modèles sont construits de sorte à représenter le degré minimum de complexité des processus physiques, chimiques et biologiques nécessaires pour quantifier la toxicité d'un composé et atteindre le pouvoir prédictif requis selon le contexte décisionnel (Margiotta-Casaluci et al., 2016). Les modèles *in silico* présentent un certain nombre de limites freinant leur acceptation comme outils indispensable pour

l'évaluation de risques. En effet, leur validation et leur développement reposent sur des données dont la disponibilité, la quantité et la qualité ne sont pas toujours de mise. De plus, leur domaine de validité est limité et défini approximativement. Finalement, la limite principale concerne le fondement prédictif et suggestif, rendant les conclusions hypothétiques. Cependant, ils offrent de nombreux avantages comparés aux tests *in vivo* et *in vitro*. Ils sont moins onéreux, hautement reproductibles et sans cesse optimisés, avec pour ambition de remplacer l'utilisation d'animaux (Erhirhie et al., 2018; Grech et al., 2017a; Leist et al., 2017).

1.1. La modélisation pharmacocinétique

La pharmacocinétique est définie par l'action d'un organisme sur un xénobiotique dont le devenir est régi par quatre processus ; l'absorption, la distribution, le métabolisme et l'excrétion (ADME). Elle permet de décrire l'évolution des concentrations d'un composé en fonction du temps selon la dose d'exposition (Grech et al., 2017a). Différents paramètres pharmacocinétiques, comme la biodisponibilité, le volume de distribution, la clairance et le temps de demi-vie, qui caractérisent la pharmacocinétique d'un xénobiotique donné, peuvent être déterminés par différentes approches. La première est l'analyse non-compartmentale permettant de déterminer les paramètres pharmacocinétiques à partir de points expérimentaux, tels que la concentration et le temps maximal, ou d'équations mathématiques simples afin de décrire les processus ADME (Gabrielsson and Weiner, 2012). La deuxième approche, dont les principes sont détaillés ci-après, modélise un organisme ou un système d'étude par un ou plusieurs compartiments, afin de quantifier les processus biologiques par une approche semi-mécaniste. Une signification physiologique et anatomique peut également être considérée pour ajouter une composante mécaniste au modèle (Dixit et al., 2003; Jusko, 2013). Grâce à la modélisation, de nouveaux paramètres pharmacocinétiques peuvent être définis et utilisés pour élaborer les relations entre le profil cinétique d'un composé et la physiologie. Par exemple, les modèles compartimentaux permettent de définir la clairance spécifique de différents processus d'élimination (Urso et al., 2002).

1.1.1. Les modèles compartimentaux classiques

Les modèles pharmacocinétiques classiques suivent une approche compartimentale considérant que le système d'étude, par exemple un organisme entier ou des cellules en culture, est réduit à un ou plusieurs compartiments pour décrire la cinétique d'un composé, comme présenté par la Figure 1.

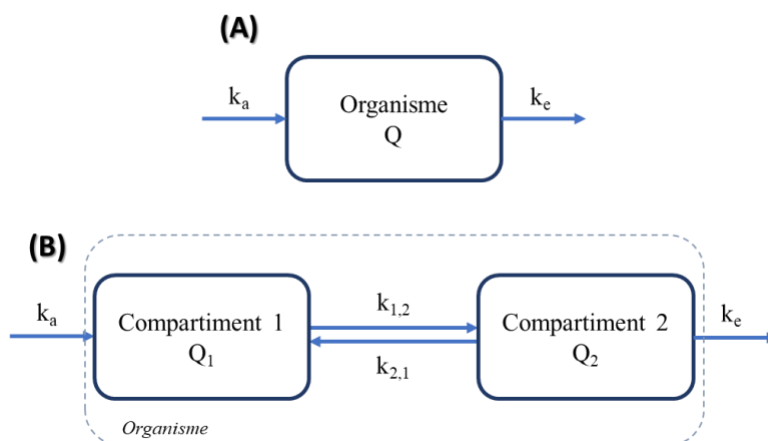


Figure 1. Schéma d'un modèle pharmacocinétique mono (A) et bi-compartmental (B), avec Q , Q_1 et Q_2 les quantités du composé dans les compartiments, k_a et k_e respectivement les constantes de vitesse d'absorption et d'élimination, $k_{1,2}$ et $k_{2,1}$ les constantes de vitesse de transfert.

Les modèles mono-compartmentaux considèrent que la concentration est homogène au sein du système. Les modèles multi-compartmentaux supposent que les concentrations peuvent différer au sein de différents compartiments caractérisant le système (Grech et al., 2017a).

Les modèles pharmacocinétiques sont définis par un système d'équations différentielles typiquement linéaire :

$$\frac{dC_i}{dt} = \sum_i (k_i \times C_i^{n_i}) \quad (1)$$

Avec dC_i/dt la variation de concentration dans le compartiment i au cours du temps, k_i une constante de vitesse et n_i un exposant déterminant l'ordre de la cinétique.

Un processus d'ordre zéro ($n = 0$) peut être dû à des phénomènes de saturation métabolique ou des systèmes de transport. La variation de concentration n'est pas dépendante de la quantité en composé restante. Un processus de premier ordre ($n = 1$) spécifie un modèle

cinétique linéaire avec des concentrations de composé dans les compartiments proportionnelles à la dose d'exposition. A faibles doses d'exposition, la majorité des composés étudiés en santé – environnement suivent approximativement cette cinétique, caractérisée par une variation de concentration proportionnelle à la quantité en composé restante et des paramètres pharmacocinétiques indépendants de la dose d'exposition (Borowy and Ashurst, 2020). Ainsi, la plupart du temps précédés d'une analyse non-compartimentale des paramètres cinétiques, les modèles compartimentaux de premier ordre utilisés en évaluation des risques permettent de quantifier les caractéristiques ADME des composés dont les processus toxiques restent à établir, suivant différentes voies d'exposition (Mackay and Fraser, 2000; Scholz et al., 2013a; Shin et al., 2019).

1.1.2. Les modèles physiologiques

Au cours des années 1970-1980, la modélisation pharmacocinétique basée sur la physiologie (PBPK) a été développée pour simuler les concentrations internes des toxiques, reconnues plus prédictives des réponses biologiques que les doses d'exposition. Les modèles PBPK considèrent l'anatomie, la physiologie, la biochimie et la physicochimie de l'organisme étudié pour simuler le devenir d'une substance dans un organisme. Ils mettent en évidence les organes ayant une action sur la substance (métabolisme, transport, dégradation, stockage...), regroupent ceux ayant peu d'importance sur les processus ou possédant grossièrement des caractéristiques similaires en un seul compartiment, puis les relient par la circulation sanguine artérielle et veineuse si l'organisme d'étude en est pourvue, comme présenté par la Figure 2 (Brochot et al., 2014).

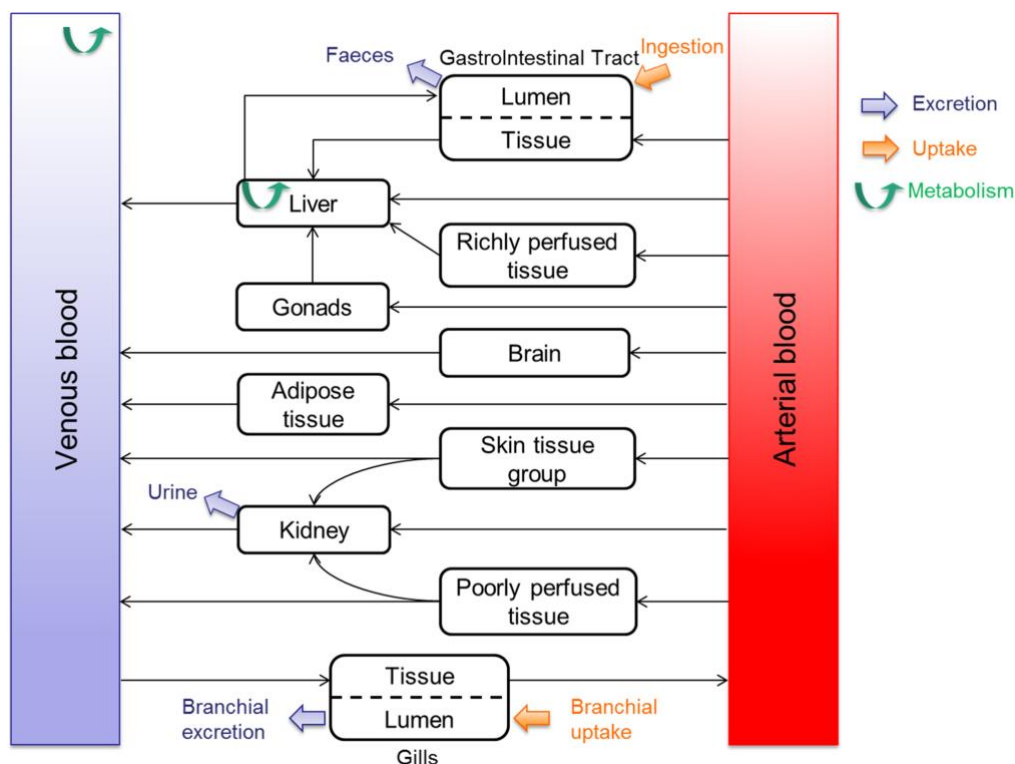


Figure 2. Modèle pharmacocinétique basé sur la physiologie du poisson adulte. L'absorption, l'élimination et le métabolisme sont représentés par les flèches orange, bleue et verte, respectivement (Grech et al., 2019).

Les échanges entre les organes, matérialisés par des compartiments, se basent sur des transports de masse, retranscrits en équations différentielles. La modélisation PBPK permet de décrire quantitativement les processus ADME des composés en corrélant les déterminants critiques que sont les volumes tissulaires, les débits physiologiques, les taux d'absorption, la diffusion membranaire, les taux et coefficients de partage tissu – sang et les réactions biochimiques (Krishnan and Peyret, 2009; Munir et al., 2018; Pearce et al., 2017). Ils peuvent également considérer des facteurs impactant ces déterminants, tels que l'âge ou les conditions d'exposition à un composé, pour étudier des populations spécifiques (Beaudouin et al., 2010; Lipscomb et al., 2012). En évaluation des risques, la modélisation PBPK est de plus en plus performante et concerne divers champs d'applications. Elle se revendique pertinente pour analyser des données sur les mécanismes pharmacocinétiques, orienter les protocoles d'études expérimentales, caractériser la variabilité et l'incertitude physiologiques et pharmacocinétiques, extrapoler les données *in vivo* et *in vitro* entre les espèces, les stades biologiques, les voies d'exposition et le temps, et finalement étayer l'évaluation globale de l'exposition à un composé (Chou and Lin, 2019; Jaroch et al., 2018; Mumtaz et al., 2012; Tan et al., 2018; WHO, 2010).

1.2. La modélisation pharmacodynamique

La pharmacodynamie est définie par l'action d'un médicament sur un système biologique. Elle permet d'étudier la relation entre la concentration interne relative à l'interaction d'un composé avec sa cible et les effets biochimiques, moléculaires et physiologiques résultants. Un effet peut être décrit par son intensité, son délai d'apparition ainsi que sa durée, et quantifié selon différentes approches : par description des effets en fonction du temps, de la dose d'exposition et par élaboration de la relation entre la concentration – effet en fonction du temps (PK/PD) (Gabrielsson and Weiner, 2006).

1.2.1. Les modèles empiriques

Les modèles de relation quantitative structure – activité (QSAR)

Le principe de la relation structure – activité repose sur la supposition que l'activité biologique des molécules est le reflet de leur structure et que des molécules similaires ont des activités similaires. Les modèles de relation quantitative structure – activité (QSAR) permettent de prédire les propriétés physico-chimiques, biologiques et le devenir dans l'environnement des composés à partir de leur structure moléculaire (Lomba et al., 2019; Papa et al., 2014). Historiquement, l'un des premiers modèles QSAR a été développé en 1962 par (Hansch et al., 1962) pour estimer la concentration chimique (C) en fonction du coefficient de partage octanol – eau (π) et la constante de Hammett (σ) relative aux effets électroniques d'un substituant (Hammett, 1937), selon l'équation suivante (Raies and Bajic, 2016) :

$$\log\left(\frac{1}{C}\right) = 4,08\pi - 2,14\pi^2 + 2,78\sigma + 3,36 \quad (2)$$

L'objectif de la modélisation QSAR est de déterminer les facteurs impactant l'activité mesurée d'un composé, toxique ou non, pour un système biologique particulier, et d'avoir un aperçu du mécanisme et du comportement du système étudié (Kar et al., 2018). Les méthodes de modélisation utilisées sont de plusieurs types : corrélation, régression et classification. En toxicologie, les analyses de régression sont les plus courantes car elles apportent un caractère quantitatif et une base mécaniste pour la toxicité modélisée contrairement aux techniques de classifications qualitatives (Schultz et al., 2003).

En évaluation des risques, les modèles QSAR permettent de fournir des informations sur les dangers des composés en réduisant le temps, le coût et le recours aux animaux pour les tests de toxicité. Cependant, peu de modèles mécanistes existent (Benigni et al., 2007; Ringeissen et al., 2011) et leur nature empirique réduit leur application (par les organismes de réglementation) à la hiérarchisation des composés existants et nouveaux pour une évaluation ultérieure de leur toxicité (Russom et al., 2003; Schmieder et al., 2003; Walker et al., 2003). De plus, les modèles QSAR ne prédisent l'activité que d'une famille ou classe de composés alors que la réglementation, par exemple des pesticides et produits biocides, peut demander l'évaluation de formulations complexes (Scholz et al., 2013a). Dans l'optique scientifique actuelle de proposer des méthodes alternatives aux expérimentations animales, l'Organisation de Coopération et de Développement Economiques (OCDE) propose une « boîte à outils » QSAR pour rendre la technologie accessible, transparente et moins exigeante en termes de coûts d'infrastructure, dans le but d'accroître l'acceptation réglementaire des méthodes QSAR. Cette « boîte à outils » est un logiciel destiné à être utilisé par les gouvernements, l'industrie chimique et d'autres parties prenantes pour combler les lacunes des données (éco)toxiques nécessaires à l'évaluation des dangers des produits chimiques (OECD, 2007).

Les modèles dose-réponse

D'après la définition de l'Organisation Mondiale de la Santé (OMS), l'évaluation de la dose-réponse permet de déterminer la relation quantitative entre des doses d'exposition ou internes et les réponses biologiques. La réponse peut être une incidence mesurée, observée ou une modification du niveau de réponse, un pourcentage de réponse dans des groupes de sujets ou une population, la probabilité d'occurrence ou de changement du niveau de réponse au sein d'une population (WHO, 2010). En évaluation des risques, simultanément à l'analyse descriptive des données, les modèles empiriques dose-réponse sont utilisés pour quantifier statistiquement la relation entre la dose d'exposition d'un composé et les réponses biologiques ou toxicologiques au sein de populations cibles, en considérant la variabilité et l'incertitude (Mushak, 2013). Les plus couramment rencontrés pour prédire une probabilité de réponse suite à l'exposition d'un composé sont les modèles de Weibull, probit et logit. Ils modélisent une courbe dose-réponse dont l'équation, par dérivation, permet d'établir les concentrations effectives (EC) responsables de 10 ou 50% de variation de la réponse initiale (Christensen et al., 2009; Christensen and Chen, 1985).

1.2.2. Les modèles mécanistes

Les modèles PBPK/PD

Au cours des dernières années, les analyses PK/PD initialement basées sur l'élaboration de la relation dose-réponse se sont transformées en des modèles mathématiques plus sophistiqués pour comprendre le mécanisme d'action des composés. Cette évolution résulte principalement de l'amélioration des méthodes analytiques, des matériels informatiques et des logiciels, de l'intérêt réglementaire et académique accru et du raffinement continu des modèles pharmacodynamiques basés sur des mécanismes physiologiques (PBPK/PD) (Csajka and Verotta, 2006). En évaluation des risques, les premières études menées pour l'identification des dangers facilitent la détermination des organes, des tissus et des systèmes affectés par l'exposition à un composé. Pour l'évaluation dose-réponse, les données décrivant les réponses sont généralement interprétées à partir de la dose externe d'exposition (Lipscomb et al., 2012).

Les modèles PBPK/PD proposent, par couplage avec un modèle PBPK, une approche mécaniste pour quantifier l'effet pharmacodynamique d'un composé par intégration de l'évolution des concentrations externes et internes au cours du temps. Ils permettent de décrire simultanément les processus ADME et l'effet des composés à différents niveaux biologiques, cellulaire ou tissulaire, par représentation mécaniste et physiologique des interactions dans le système biologique d'étude, et ainsi de comprendre les relations entre concentration d'exposition, temps et effet (Ducrot et al., 2016; Kuepfer et al., 2016), comme présenté par la Figure 3, illustrant le modèle PBPK/PD développé par (Sharma et al., 2017) pour étudier la neurotoxicité de l'acide perfluorooctanesulfonique (PFOS) chez l'Homme.

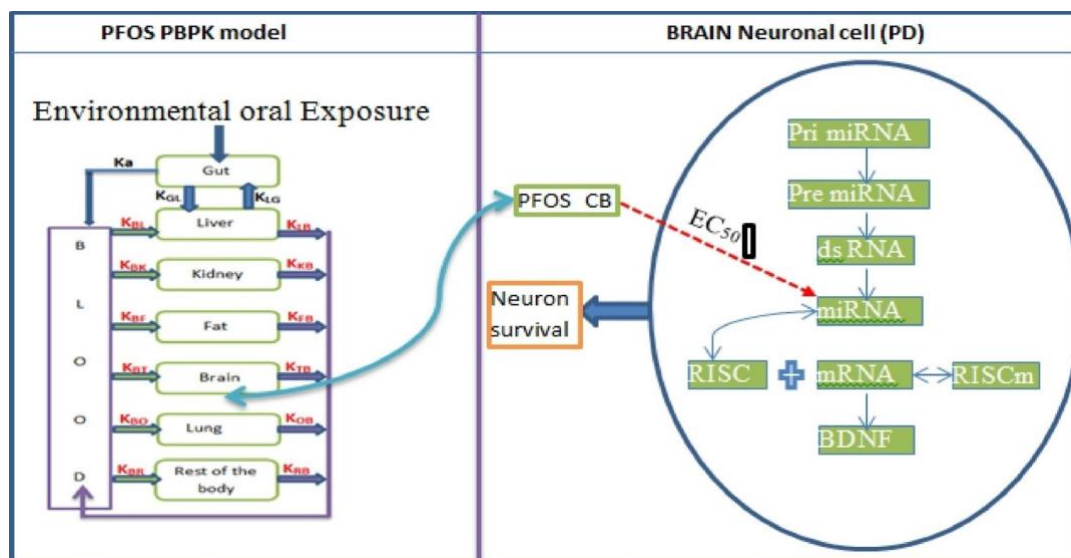


Figure 3. Schéma d'un modèle PBPK/PD de l'acide perfluorooctanesulfonique (Sharma et al., 2017).

Les processus ADME peuvent varier entre les espèces animales et l'Homme. Appréhender et déterminer les différences inter-espèces est essentiel pour extrapoler de façon pertinente et argumentée les données de réponses biologiques animales à l'humain (Lipscomb et al., 2012). Par ses considérations mécanistes, la modélisation PBPK/PD est un outil utile en évaluation des risques pour extrapoler les concentrations toxiques entre différentes espèces, mais également entre les études *in vitro* et *in vivo*, les études précliniques et cliniques, de la dose externe d'exposition à la dose systémique, pour relier la physiologie et la génétique à la disposition des composés dans les populations animales et humaines, et enfin, pour simuler la temporalité des processus toxiques (Ashauer and Escher, 2010; Dixit et al., 2003).

Les modèles de toxicologie systémique

Les voies de signalisation cellulaires ou métaboliques expliquent comment les gènes et protéines interagissent entre eux pour déclencher les fonctionnalités cellulaires ainsi que la génération de molécules ou de métabolites. La Figure 4 présente schématiquement la cascade cellulaire impliquée dans le métabolisme du glutathion, acteur essentiel des réactions de détoxification et de défense contre les espèces réactives de l'oxygène (Reed et al., 2008). La production de données génomiques a permis d'accroître les connaissances sur les mécanismes génétiques d'un grand nombre de processus toxiques. De plus, l'imagerie cellulaire permet la compréhension de la corrélation entre les paramètres, les gènes, les protéines et les voies pour

guider l'intégration de la dynamique moléculaire et cellulaire à l'aide d'approches de modélisation des systèmes (Sipes et al., 2011a).

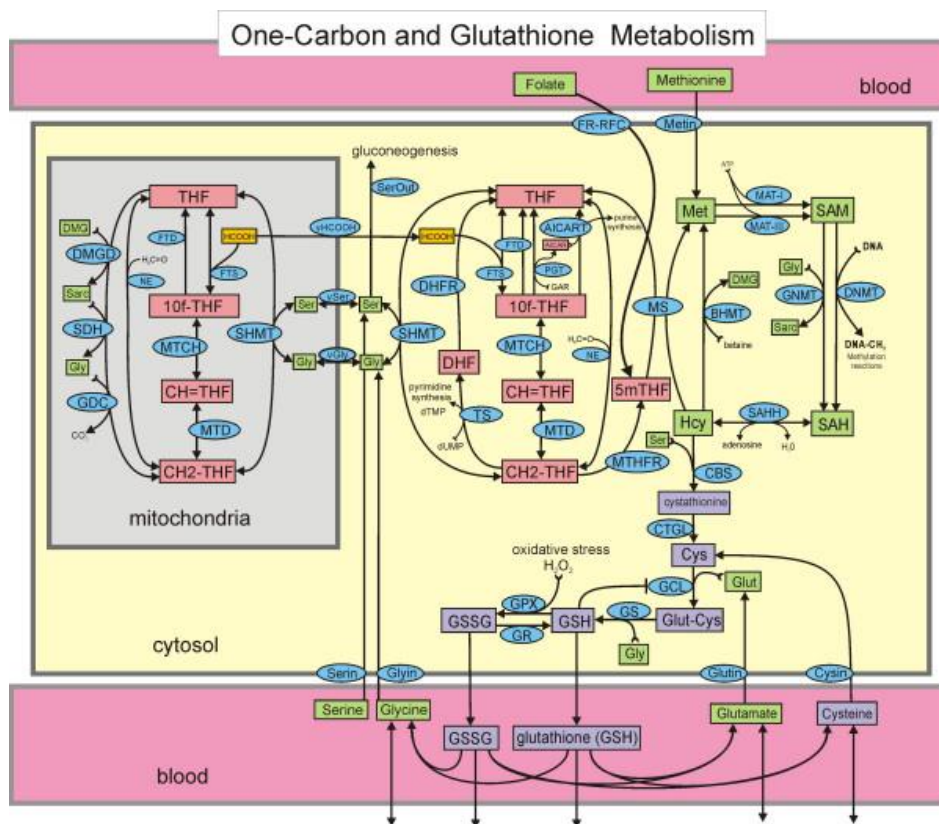


Figure 4. Schéma du modèle de la voie du métabolisme du GSH par Reed et al. (2008)

Des modèles mathématiques de type différentiel, ou discret (Booléen par exemple), parfois stochastiques, de l'activité de ces voies décrivent le comportement des cellules pour étudier les mécanismes moléculaires d'action selon des conditions d'exposition particulières et prédire les conséquences potentielles des perturbations (mutations ou changements de l'expression) des protéines *etc.* Cette approche holistique offre une alternative aux approches centrées uniquement sur les gènes pour contribuer à l'interprétation de la variabilité génomique des phénomènes toxicologiques complexes. En d'autres termes, cela signifie que l'exposition à des composés peut conduire à des variations dans la séquence d'ADN (Amadoz et al., 2015; Peña-Chilet et al., 2019; Rahman et al., 2018).

1.2.3. Les modèles intermédiaires entre empiriques et mécanistes

Les modèles quantitatifs de chemins d'effets néfastes (qAOP)

Les chemins d'effets néfastes (AOP : adverse outcome pathways) modélisent le processus toxique d'un composé en reliant un événement initiateur moléculaire (MIE : molecular initiating event) à l'effet toxique observé (AO : adverse outcome) par l'intermédiaire d'une chaîne d'événements clés (KE : key event) à des niveaux croissants d'organisation biologique ; des réponses cellulaires jusqu'à la dynamique des populations, selon la Figure 5. Construits à partir d'informations mécanistes obtenues *in vitro*, ils sont axés sur l'identification des cibles moléculaires potentielles d'un composé et les effets *in vivo* à différents niveaux biologiques qui en résultent.

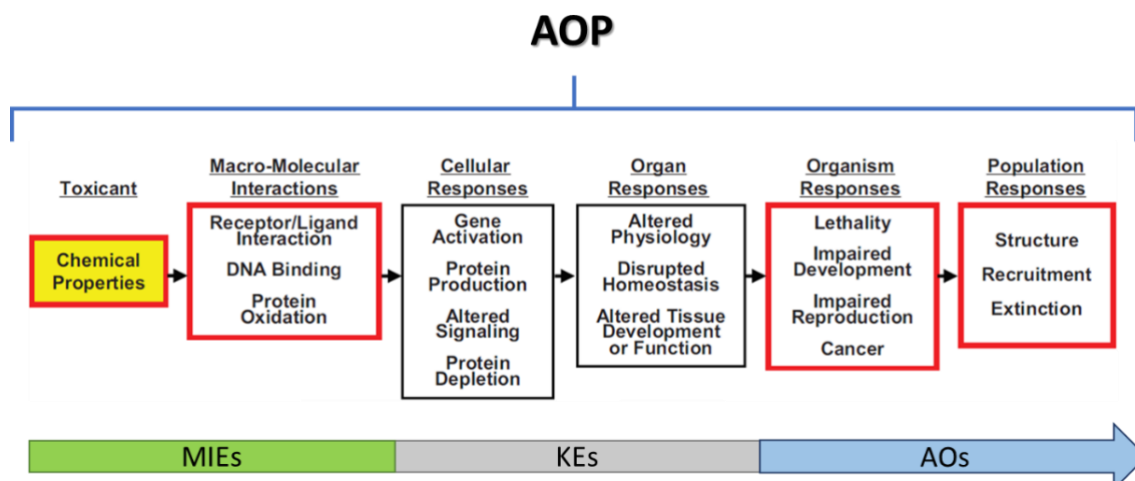


Figure 5. Représentation schématique des différents niveaux organisationnels d'un AOP, adapté à partir de (Ankley et al., 2010).

L'intégration de la pharmacocinétique et de la pharmacodynamie ajoute à la nature fondamentalement qualitative des AOP une approche quantitative (qAOP) pour la prédiction de la toxicité des composés. La considération de tous les niveaux d'organisation biologique permet d'évaluer le risque sur la partie du réseau (niveau moléculaire ou populationnel) pour différentes espèces et sur la voie d'exposition, rendant ainsi pertinent l'utilisation de qAOP pour la gestion des risques et la prise de décisions en santé – environnement (Leist et al., 2017; Margiotto-Casaluci et al., 2016), selon la Figure 6. En fonction des données disponibles, les AOP peuvent être élaborés à partir de modèles statistiques, bayésiens, de régression, d'équations différentielles ordinaires (ODE) et individuels (Groh et al., 2015; Spinu et al.,

2020), afin de répondre aux questions posées au cours des quatre étapes constitutives de l'évaluation des risques : évaluation des dangers, évaluation de la relation dose-réponse, évaluation de l'exposition et caractérisation des risques (Perkins et al., 2019).

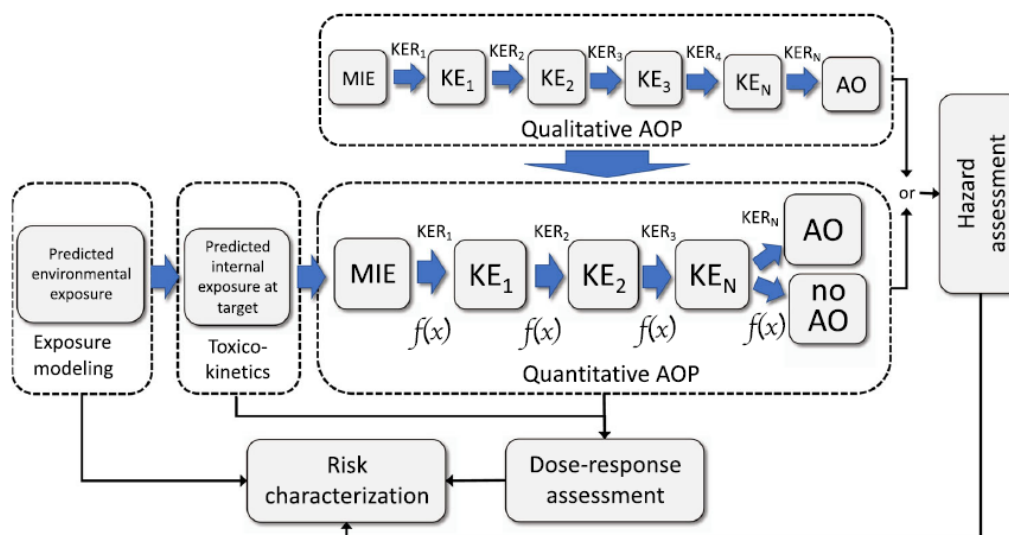


Figure 6. Organisation et utilisation des chemins de l'effet néfaste (AOP) pour l'évaluation des risques (Perkins et al., 2019).

Les modèles de survie

Décrire l'apparition de phénomènes dans le temps pour déterminer les facteurs causaux représente une composante essentielle de l'évaluation des risques. Dans le domaine santé – environnement, l'occurrence des effets observés suite à l'exposition à un composé est relevée à des temps précis, prédéterminés en fonction du protocole de l'étude. Par exemple, pour les tests de toxicité animale, l'OECD recommande d'analyser séparément les doses-réponses aux différentes concentrations d'exposition étudiées, pour chaque temps d'observation fixé (OECD, 2013). Cependant, les réponses, appelées évènements en analyse de survie, peuvent se produire à n'importe quel instant au cours de l'étude, biaisant les temps d'occurrence. Ces processus d'observation mènent à des données censurées, pouvant être de trois types :

- Censure à droite : les évènements se produisent après le temps de la dernière observation ;
- Censure à gauche : les évènements se produisent avant le temps de la première observation ;
- Censure par intervalle : les évènements se produisent entre deux temps d'observation (Moore, 2016; Philippe Saint Pierre, 2015).

Dans certaines études, une partie des individus initialement inclus n'est pas observable et les analyses se font à partir de sous-échantillons. Il s'agit d'un processus de troncation dont trois types sont également possibles : troncations à droite, à gauche ou par intervalle (Saint Pierre, 2015).

Les évènements sont en partie conditionnés par les phénomènes passés. Séparer les observations par groupes de temps réduit donc le nombre de données par analyse et ignore cette dépendance temporelle, entraînant une perte d'information importante. Les modèles de survie ont été élaborés afin de répondre à cette problématique en proposant de modéliser la distribution du temps d'apparition d'un évènement, à partir d'un temps initial fixé et de données observées chez le système biologique d'étude. Ce sont donc des modèles stochastiques car ils prédisent des probabilités d'occurrence. Ils permettent de décrire des données univariées (par exemple, l'âge du décès) ou multivariées (dose d'exposition et effet) (Bradburn et al., 2003a, 2003b; Clark et al., 2003a, 2003b; Jiang and Fine, 2007; Lee and Go, 1997; Leung et al., 1997; Mahnken et al., 2008). Ils sont communément utilisés au cours des études cliniques et des analyses de survie (Gould et al., 2015; Karmen et al., 2019). Le modèle de survie le plus simple décrit le passage de l'état 1 « vivant », à l'état 2 « mort », comme illustré par la Figure 7. Le taux de transition $q_{1,2}$ peut relier l'exposition instantanée aux médicaments ou aux substances toxiques et l'évolution temporelle de la mortalité observée chez les individus exposés. Il est à noter qu'un état est par définition statique, à la différence de l'évènement, dynamique au cours du temps. Un individu change d'état à la suite de l'occurrence d'un évènement.

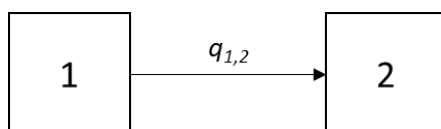


Figure 7. Schéma du processus basique d'analyse de survie.

Les analyses de survie font l'hypothèse que le temps de survie T (temps passé dans l'état 1) est une variable aléatoire. La fonction de densité de probabilité, $f(t)$, et la fonction de distribution cumulative, $F(t)$, caractérisent la distribution de cette variable aléatoire. La fonction de survie est définie par :

$$S(t) = 1 - F(t) = P(T > t), \quad t \geq 0 \quad (3)$$

La probabilité de passer d'un état initial i ($X(0)$) à un état aléatoire j ($X(t)$) au cours du temps t est définie comme suit :

$$p_{i,j}(t) = P\{X(t) = j | X(0) = i\} \quad (4)$$

Les modèles multi-états

Les modèles multi-états sont des modèles de survie plus complexes, considérant que plusieurs évènements peuvent découler de l'état initial d'observation, comme présenté par la Figure 8. Dans ce cas, il est possible que les observations soient indirectement liées, puisque si la transition de l'état 1 à l'état 2 se produit, l'individu ne pourra plus se trouver dans l'état 3. Cela conduit à un biais d'observation de l'état 3. Par exemple, des embryons observés morts ne développent pas de malformations cardiaques ultérieures. Ainsi, pour des concentrations létales, la prévalence de malformations cardiaques diminue car les embryons meurent avant le développement du cœur. Il s'agit d'un processus des risques concurrents. Pour résoudre ce genre de problèmes au cours des analyses de données de survie, les modèles multi-états markoviens peuvent décrire la distribution des temps de séjour des sujets dans les différents états, avec des probabilités d'évènement dans un état donné non dépendantes de la façon dont cet état a été atteint (Andersen et al., 2012; Beyersmann et al., 2012; Meira-Machado et al., 2009).

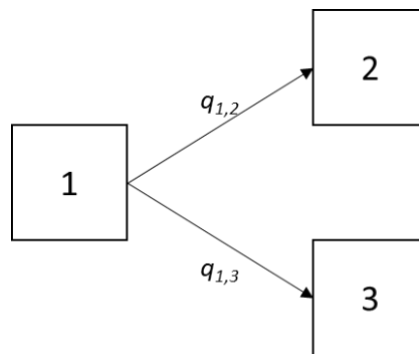


Figure 8. Schéma d'un processus simple des risques concurrents.

Les modèles multi-états markoviens sont très généraux, et peuvent considérer des évènements intermédiaires ainsi que des retours possibles entre états (Farewell and Tom, 2014; Putter et al., 2007), comme illustré par la Figure 9.

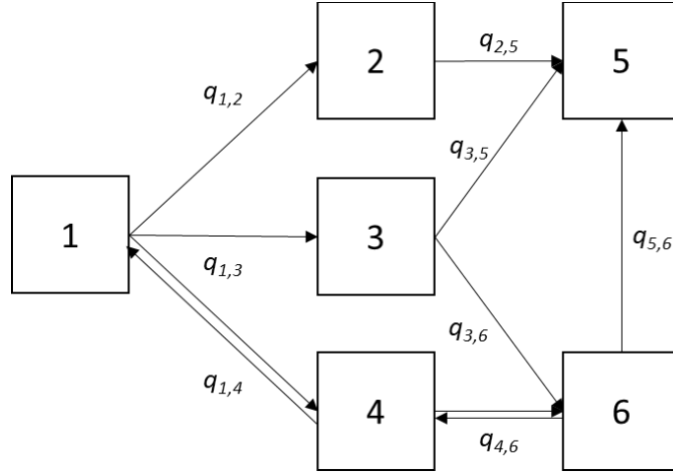


Figure 9. Schéma d'un modèle multi-états, avec évènements intermédiaires et possibilité de retour entre états.

Comme dans les modèles simples de survie, une série de taux de transitions $q_{i,j}(t, z(t))$ détermine le risque de passer de l'état i à l'état j , en fonction du temps et de variables individuelles $z(t)$, pouvant être également dépendante du temps (Jackson, 2011).

Deux simplifications sont possibles à partir du modèle général markovien :

- Les modèles multi-états markoviens homogènes dans le temps considèrent que les taux de transition dépendent seulement de l'état initial et sont indépendants du temps (Putter et al., 2007; Zare et al., 2014) :

$$q_{ij}(t, z(t)) = q_{ij}(z) \quad (5)$$

- Les modèles semi-markoviens considèrent que les taux de transition ne dépendent que de l'état initial i et du temps d'entrée t_j dans l'état suivant j (Meira-Machado et al., 2009) :

$$q_{ij}(t, z(t)) = q_{ij}(t - t_j, z) \quad (6)$$

Le modèle peut également être étendu aux processus de Markov à état cachés si l'appartenance aux états ne peut être donnée directement par les observations (Lavielle, 2018).

1.3. L'inférence statistique sur les paramètres des modèles

1.3.1. L'inférence bayésienne

On entend généralement par inférence statistique l'estimation des valeurs des paramètres d'un modèle mathématique sur la base d'un ensemble d'observations (« données »). Il y a actuellement deux grandes classes de méthodes d'inférence statistique : les approches bayésiennes et les approches fréquentistes. L'inférence bayésienne considère les paramètres à estimer comme des variables aléatoires et les données comme des valeurs fixes, alors que l'inférence fréquentiste fait juste le contraire.

L'inférence fréquentiste se fonde généralement sur la considération d'ensembles hypothétiques de données répliquées. Traditionnellement, les méthodes et outils fréquentistes comprennent des tests d'hypothèse, des intervalles de confiance et des estimateurs de paramètres. Une caractéristique fondamentale de l'inférence fréquentiste est l'évaluation des probabilités d'erreur, *i.e.* des probabilités de mauvaises décisions (Gelman and Hennig, 2017).

L'avantage de l'approche bayésienne est d'intégrer facilement des connaissances *a priori*, (disponibles indépendamment du jeu de données en cours d'analyse) ou des données obtenues séquentiellement (Gelman and Hennig, 2017). Cependant, l'approche bayésienne requiert de définir une distribution *a priori* des paramètres à estimer ; ce choix n'est pas toujours facile, et peut être perçu comme subjectif, d'où certaines controverses et critiques sur le choix le plus approprié en l'absence de connaissances *a priori* (O'Hagan, 2008). Le théorème de Bayes permet de dériver la distribution *a posteriori* de paramètres θ en fonction d'un *a priori* et de nouvelles données selon la formule suivante :

$$f(\theta|y) \propto f(\theta) \times f(y|\theta) \quad (7)$$

Avec $f(\theta|y)$ la densité *a posteriori* de θ , $f(\theta)$ la densité *a priori* et $f(y|\theta)$ la distribution de probabilité des données, conditionnellement à θ . Cette dernière distribution est la seule considérée dans l'approche fréquentiste, elle est aussi appelée fonction de vraisemblance. La densité *a posteriori* a par définition une intégrale égale à 1. Cependant, le produit de la densité *a priori* par la vraisemblance n'a généralement pas une intégrale égale à 1. La constante de proportionnalité entre ce produit et la densité *a posteriori* est égale à la probabilité totale des données (intégrale du produit par rapport à θ). Cette constante n'est en général pas connue, sauf

dans des cas relativement simples (O'Hagan, 2008). Ceci a limité pendant longtemps l'application des méthodes bayésiennes.

Dans l'inférence bayésienne, toutes les décisions à prendre sont dérivées de la distribution *a posteriori*, puisqu'elle contient toute l'information disponible sur les paramètres après observation des données. Divers critères de décision peuvent être appliqués pour obtenir une estimation ponctuelle de chaque paramètre θ .

1.3.2. Les modèles multi-niveaux

Dans les situations de recherche les plus courantes, l'évaluation directe de l'ensemble d'une population d'intérêt n'est pas possible. Ainsi, les études expérimentales se portent généralement sur un échantillon pour déterminer les paramètres de la population générale. En évaluation des risques, il est important d'évaluer l'incertitude et la variabilité affectant les résultats expérimentaux afin d'accorder confiance aux prédictions des modèles. L'incertitude est principalement le résultat d'un manque de connaissance nourri par la précision des instruments de mesure, l'utilisation de données pour ajuster les modèles PK et PD causant du « bruit » sur les valeurs des paramètres résultantes, le réductionnisme de l'échantillon comparé à l'étude d'une population entière, les simplifications des modèles et les extrapolations *in vitro* – *in vivo* ou inter-espèces. La variabilité considère les différences de caractéristiques épigénétiques, génétiques, environnementales et leurs interactions, entre les individus (variabilité interindividuelle) ou dans le temps pour un individu donné (variabilité intra-individuelle). Si l'incertitude peut être réduite par enrichissement des connaissances des processus toxiques par exemple, la variabilité est inhérente aux populations. Ainsi, pour considérer l'incertitude impactant l'évaluation de la variabilité, des modèles statistiques de populations ont été développés, initialement pour les études pharmacocinétiques puis élargis aux modèles dose-réponse. L'objectif des modèles de population est d'extraire, à partir de données sur des unités individuelles (par exemple les lots d'embryons de poisson zèbre spécifiques à chaque temps d'observation du protocole expérimental de l'OCDE), une description quantitative de la variabilité de la cinétique d'un composé (PK) ou de l'occurrence d'un effet (PD) au sein d'une population. Conceptuellement, le même modèle décrit les données pour chaque unité et ses paramètres varient aléatoirement entre les unités ou les occasions d'observation d'une même unité. La variabilité, n'ayant pas de fondement mécaniste, est décrite par une distribution probabiliste. Par exemple, dans le modèle PK-PD multi-états développé

pour les travaux de cette thèse, le taux d'éclosion des embryons se distribue log-normalement autour d'une moyenne de population, avec une variance de population relative à la variabilité populationnelle. Les moyennes et les variances de la population sont appelées paramètres de population. Ils sont estimés par ajustement des données, avec les paramètres du modèle pour chaque unité. D'autres paramètres de population peuvent être définis, comme par exemple les covariances des paramètres si certains paramètres semblent corrélés au niveau populationnel (Bois, 2001; Wakefield and Rahman, 2000).

1.3.3. La calibration des modèles développés pour les travaux de thèse

Les modèles développés pour les travaux de cette thèse suivent des méthodes de Monte Carlo par chaînes de Markov (MCMC) pour calculer des approximations numériques d'intégrales multidimensionnelles, selon une approche bayésienne. Dans les statistiques bayésiennes, le développement récent des méthodes Monte Carlo avec chaînes de Markov simulées – par exemple, l'échantillonnage de Gibbs (Jones et al., 2017) – permet de calculer des modèles multi-niveaux avec de nombreux paramètres inconnus. Ces méthodes créent des échantillons d'une variable aléatoire continue multidimensionnelle, avec une densité de probabilité proportionnelle à une fonction de densité connue. Plusieurs chaînes sont typiquement générées à partir d'un ensemble de points choisis arbitrairement et suffisamment éloignés les uns des autres (Banerjee et al., 2015). Dans un cadre bayésien, la densité cible est la distribution *a posteriori* des paramètres du modèle, étant donné le modèle statistique, les distributions de paramètres *a priori* et les données pour lesquelles une fonction de vraisemblance peut être calculée (Bois, 2009; Smith and Roberts, 1993).

Les modèles pharmacocinétiques et pharmacodynamiques suivent des approches fréquentistes ou bayésiennes. Pour les modèle PBPK, l'approche bayésienne est plus appropriée puisque leurs paramètres ont en général une signification physique directe et bénéficient d'estimation *a priori* indépendantes des données à analyser. C'est moins souvent le cas pour les modèles multi-états, également moins utilisés et moins spécialisés. Cependant, même pour ces modèles l'analyse bayésienne est de plus en plus proposée (Williams et al., 2020). Nous adopterons cette approche car elle rend assez facile l'application des modèles multi-niveaux.

1.4. Les intérêts et limites de la modélisation pour l'extrapolation à l'Homme

Pour extrapoler les données de toxicité animale et *in vitro* d'un composé à l'Homme, il est primordial de déterminer sa pharmacocinétique, relative à son comportement au sein du corps humain, et sa pharmacodynamie, relative à ses effets au niveau des cibles biologiques. Malgré une distinction courante des deux processus, ils forment un continuum puisque l'occurrence et l'intensité des effets dépendent de l'évolution de la concentration du composé au cours du temps, au sein d'un système d'étude. En raison des effets pharmacocinétiques, les sensibilités tissulaires et organiques observées chez l'animal ou *in vitro* peuvent être différentes chez l'Homme. Ainsi, la première étape pour extrapoler de façon pertinente des données de toxicité est de quantifier la pharmacocinétique des composés. Pour ce faire, les modèles mécanistes PBPK représentent actuellement la meilleure alternative pour extrapoler les données à l'Homme, avec un large domaine d'applicabilité et permettre de prédire la pharmacocinétique d'un composé au-delà du temps d'observation et des doses testées. La Figure 10 illustre les principales données essentielles pour informer sur la structure des modèles et les valeurs des paramètres. La deuxième étape consiste à prédire la réponse pharmacodynamique ou toxique chez l'Homme. Comme une même concentration d'exposition *in vitro* ou chez l'animal peut conduire à des effets différents chez l'Homme, il est important de développer des modèles mécanistes des réponses cellulaires caractérisés par une structure hiérarchique (Nadia Quignot et al., 2014).

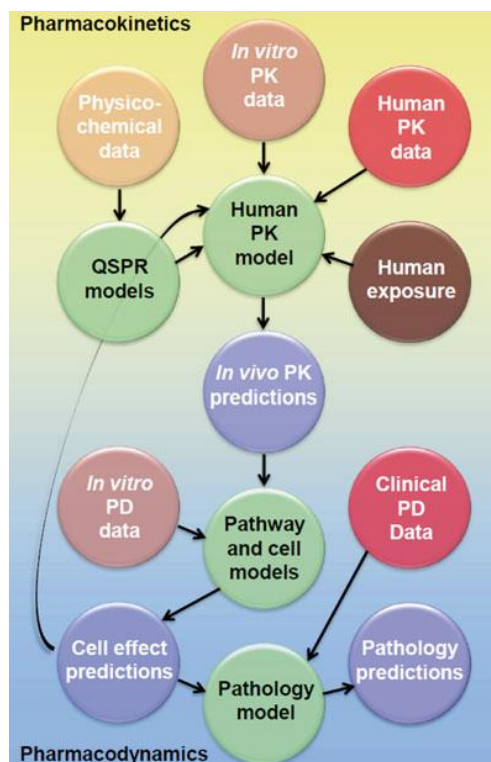


Figure 10. Principales données nécessaires à l'extrapolation quantitative des données à l'Homme et de la prédiction de la toxicité. Les prédictions (en bleu) de la pharmacocinétique *in vivo*, des effets cellulaires *in vivo*, ainsi que des paramètres pathologiques (toxicité), sont obtenues par différents modèles (en vert) capables d'intégrer des données (en rouge) provenant de diverses sources (Nadia Quignot et al., 2014).

Prédire la toxicité chez l'Homme implique également de considérer la variabilité des caractéristiques physiologiques, du métabolisme, etc. au sein de la population humaine. Si l'impact de la PK sur la PD est considéré, l'inverse l'est rarement. Cependant, l'action de la PD sur la PK est susceptible de se produire et ce avec une certaine latence et des mécanismes complexes. Les modèles prédictifs de tels effets, notamment des effets cumulatifs, de la toxicité développementale ou de la carcinogenèse, permettraient de couvrir le continuum PK – PD (Nadia Quignot et al., 2014).

Les modèles mathématiques représentent un réel atout pour valoriser les expérimentations animales et les modèles *in vitro*. Ils fournissent des informations complémentaires fondamentales pour permettre une juste extrapolation des données de toxicité des composés à l'Homme, par modélisation des processus toxiques à tout niveau de considération biologique. Cependant, le critère prédictif et représentatif caractérisant les modèles *in silico* conduit à des limites à considérer lors de l'interprétation de résultats.

Les modèles mathématiques sont validés et construits à partir de données de nature différente. Ainsi, plus le nombre de données disponibles est important et plus le modèle aura de chance d'être correct et prédictif. En évaluation des risques, le manque de données sur la toxicité des composés affecte la performance des modèles développés. Leur construction parfois trop simpliste peut ne pas suffire à modéliser justement les variables incluses, et ne permet pas d'envisager et de décrire correctement l'impact de tous les facteurs intervenant dans les processus toxiques. Finalement, malgré la volonté scientifique d'améliorer et de développer les modèles mathématiques, soutenue par des organismes mobilisés pour l'évaluation du risque humain, tels que l'OCDE, leur exploitation et utilisation sont seulement considérées par les instances règlementaires comme un support des tests *in vitro* et *in vivo* obligatoires et non un prérequis (Erhirhie et al., 2018). Suite à l'analyse de l'application des différents types de modèles d'effet, des recommandations générales pour l'utilisation de la modélisation dans l'évaluation des risques des pesticides ont été formulées mais leur intégration dans les systèmes d'évaluations n'est pas définie (Hommen et al., 2016).

2. Les données de toxicologie développementale et reproductive

2.1. Les données *in vivo*

L'utilisation des animaux pour étudier la pharmacologie et la toxicologie remonte au 19^e siècle. Initialement élaborée pour prévenir la toxicité aiguë et ainsi classer les produits chimiques, elle est à ce jour un prérequis dans divers domaines scientifiques malgré la tendance à promouvoir des méthodes alternatives prédictives (Knudsen et al., 2019).

Les expérimentations animales contribuent grandement à la compréhension des mécanismes des maladies et des comportements des toxiques au sein d'un organisme vivant (van der Worp et al., 2010). Considérer un organisme dans son entier pour étudier la toxicocinétique d'un ou plusieurs composés en mélange permet de déterminer les mécanismes d'action mis en jeu dans les processus toxiques et d'appréhender la dynamique des paramètres physiologiques impactés par le comportement d'un composé (Alzualde et al., 2018; Martinez et al., 2018; Zoupa et al., 2020). Les données animales peuvent être extrapolées pour déterminer la pharmacocinétique chez l'Homme. Par exemple, au cours des essais précliniques d'un médicament, les tests toxicologiques sont principalement effectués sur des rongeurs, des lapins, des chiens, des primates non humains et des cochons. Des échantillons sanguins sont ensuite

analysés à différents temps pour déterminer les paramètres pharmacocinétiques du composé, relatifs à l'absorption, la distribution, le métabolisme et l'excrétion (Parasuraman, 2011). Cependant, depuis quelques dizaines d'années, l'expérimentation animale connaît un réel tournant éthique. En 1959, William Russell et Rex Burch, deux scientifiques anglais, ont posé dans le livre « *The Principles of Humane Experimental Technique* » le principe des 3R (Replacement, Reduction, Refinement) visant à substituer, réduire et améliorer les expérimentations animales (Kirk, 2018). Depuis, de nombreuses réglementations et initiatives internationales ont émergé pour encadrer et éviter les dérives abusives, mais surtout pour proposer des alternatives à l'utilisation des animaux. L'OCDE a élaboré des lignes directrices et incite les scientifiques à développer des outils innovants comme des Approches Intégrées pour Tester et Evaluer (IATA) les composés chimiques (Casati, 2018; Knudsen et al., 2019). En 2007, le règlement européen « REACH », concernant l'enregistrement, l'évaluation, l'autorisation et la restriction des produits chimiques, a été adopté pour protéger harmonieusement la santé humaine et l'environnement des risques liés à leur exposition. Deux articles font en particulier référence aux expérimentations animales en stipulant leur utilisation en dernier recours et promouvant les méthodes alternatives (Sobanska et al., 2018; Taylor, 2018).

2.2. Les limites des données *in vivo* pour l'extrapolation à l'Homme

Même si les expérimentations animales restent un atout majeur dont certains domaines scientifiques peuvent difficilement se passer, des précautions doivent être prises pour émettre des hypothèses ou affirmations quant aux résultats obtenus avec des tests de toxicité animale. En évaluation des risques, il ne fait aucun doute que la meilleure espèce pour étudier la toxicité humaine est l'Homme (Erhirhie et al., 2018). Malgré l'importante avancée des techniques et outils d'analyse, l'extrapolation exacte des données animales à l'Homme ne peut être garantie en raison des variations inter-espèces anatomiques, physiologiques et biochimiques. Certaines études montrent notamment une faible reproductibilité et comparabilité des processus toxiques entre les animaux et l'Homme (Van Norman, 2019).

2.3. Les données *in vitro*

Les technologies *in vitro* trouvent leur intérêt dans différents domaines scientifiques, notamment pour répondre à des problématiques de toxicité suite à l'exposition des organismes

vivants aux xénobiotiques et aux facteurs de stress environnementaux (Crean et al., 2015). De plus, l'évolution des réglementations sur le bien-être des animaux utilisés à des fins scientifiques a accéléré leur développement. Un test réalisé *in vitro* est réalisé dans un environnement contrôlé, tel qu'une éprouvette ou une boîte de Pétri, hors d'un organisme vivant et implique communément des cellules isolées, des tissus ou des organes animaux ou humains. Les tests conventionnels de toxicité impliquant l'utilisation d'un grand nombre d'animaux sont progressivement remplacés ou complétés par des méthodes alternatives *in vitro* (Bellwon et al., 2015; European Chemicals Agency, 2016).

Les méthodes *in vitro* permettent d'étudier l'implication des processus biologiques fondamentaux et sont donc utiles pour aider à élucider les mécanismes de toxicité d'un composé (Bennet et al., 2017; Fry et al., 2019). Elles présentent plusieurs avantages car elles impliquent de petites configurations expérimentales requérant peu de substance d'essai, offrent un nombre élevé de répétitions et sont peu onéreuses à l'unité. De plus, comparés aux modèles *in vivo* identifiant les mécanismes toxiques, les systèmes *in vitro* sophistiqués permettent de s'affranchir des effets spécifiques aux espèces, par l'utilisation de lignées cellulaires humaines. Des modèles cellulaires pour pratiquement tous les tissus ou espèces d'animaux de laboratoire sont disponibles mais l'utilisation de lignées cellulaires humaines reste préférée en évaluation des risques humains car elles sont plus susceptibles de prédire les effets éventuels. L'innovation dans les méthodes *in vitro* vise à améliorer les techniques de co-culture, les systèmes de culture organoïdes, la miniaturisation des systèmes physiologiques, etc. (Erhirhie et al., 2018; McMullen et al., 2018; Rahmani et al., 2019; Rhomberg, 2010; Watson et al., 2017).

2.4. Les limites des données *in vitro* pour l'extrapolation à l'Homme

Les méthodes *in vitro* aident grandement à la compréhension des mécanismes de toxicité des composés et leur innovation importante accroît leur capacité de prédiction, atout indispensable en évaluation des risques pour l'humain. Cependant, plusieurs limites sont à considérer lors de l'interprétation de résultats de toxicité *in vitro*. Malgré la performance grandissante des systèmes *in vitro* pour reproduire au mieux les conditions *in vivo*, de nombreuses caractéristiques anatomiques et physiologiques ne sont pas considérées. En général, un seul type cellulaire est utilisé, ne permettant pas d'étudier l'interaction physiologique des types cellulaires ou tissus entre eux au sein d'un organisme entier. De plus et de ce fait, les systèmes *in vitro* ne sont pas toujours riches des multiples enzymes et autres

protéines présentes dans les cellules humaines *in vivo* et pouvant interférer avec le composé étudié. Dans certaines conditions expérimentales, une dégénérescence cellulaire peut être observée due à l'appauvrissement continu en nutriments, à l'accumulation de déchets et à un apport insuffisant en oxygène entraînant des conditions de culture anaérobie. Des lignées cellulaires peuvent également être contaminées par des micro-organismes et des composés de l'air environnant en raison de mauvaises conditions de stockage. Pour finir, il n'est pas à exclure la survenue de problème éthiques liés au don de tissus humains (Erhirhie et al., 2018).

2.5. Les tests sur embryons de poisson zèbre (ZET)

Les mammifères apparaissent comme les plus adéquats pour étudier la toxicité pour l'humain. Cependant, les expérimentations sont couteuses, strictement réglementées et longues. D'autres groupes d'espèces ont des caractéristiques physiologiques et génétiques semblables à l'Homme et sont donc des compléments ou alternatives pertinentes pour les tests de toxicité (Hill et al., 2005; Larisch and Goss, 2018; Lieschke and Currie, 2007). Le poisson zèbre, un vertébré, combine les bénéfices des expérimentations sur organisme entier et *in vitro* (dans son stade embryonnaire). Les tests de toxicité sur le poisson zèbre se sont donc développés dans divers domaines scientifiques, notamment pour modéliser les maladies humaines à l'aide de modifications génétiques et mieux comprendre et prédire la toxicité chez l'Homme. Cela est notamment possible grâce aux avancées technologiques et aux outils moléculaires disponibles pour analyser le poisson zèbre. La toxicité de composés variés peut être étudiée : métaux, métalloïdes naturels, composants synthétiques des produits de consommation, pesticides, médicaments, sous-produits de transformation industrielle et d'incinération des déchets *etc.* (Bambino and Chu, 2017; Lawrence, 2011). Cette partie décrit les caractéristiques qui font l'attrait de cette espèce, notamment pour étudier la toxicité développementale humaine.

2.5.1. Caractéristiques de l'espèce

Le poisson zèbre *Danio rerio* est un poisson téléostéen tropical appartenant à la famille des Cyprinidés, dont la taille moyenne adulte oscille autour de 35 mm pour un poids d'environ 300 mg (Hamilton, 1822; Lawrence, 2011). Il est principalement retrouvé dans les eaux douces alcalines (pH ~ 8) peu profondes des rivières, ruisseaux et rizières de l'Asie du sud (Inde, Birmanie, Bangladesh et Népal) (Engeszer et al., 2007; jayasimhan et al., 2017; Spence et al., 2007). Dans la nature, la robustesse du poisson zèbre lui permet de supporter des températures

comprises entre 12°C et 33°C. En laboratoire, sa température optimale de vie est de 28,5°C (Arunachalam et al., 2013; Holtzman et al., 2016).

La Figure 11 représente le cycle de vie du poisson zèbre. La fertilisation de l'œuf initie le développement à partir d'une cellule zygote reposant sur le sac vitellin. L'éclosion (jour 2-3 à 5) est principalement marquée par la perte du chorion (membrane semi-perméable l'entourant auparavant), l'indépendance de l'embryon qui se met à nager et par la fin de la morphogénèse des organes primaires (Kimmel et al., 1995a). Il devient alors un éleuthéro-embryon ou pro-larve. Entre le quatrième et cinquième jour, la consommation totale du sac vitellin contenant les nutriments essentiels au développement, signe son indépendance à se nourrir et définit le passage au stade larvaire (Braunbeck et al., 2015). A la fin du premier mois de développement, le poisson entre dans le stade juvénile s'étendant jusqu'au troisième mois post fertilisation. Le poisson zèbre atteint sa maturité sexuelle après 3 mois de croissance et a une durée de vie d'environ cinq ans (D'Costa and Shepherd, 2009; Stewart et al., 2014).

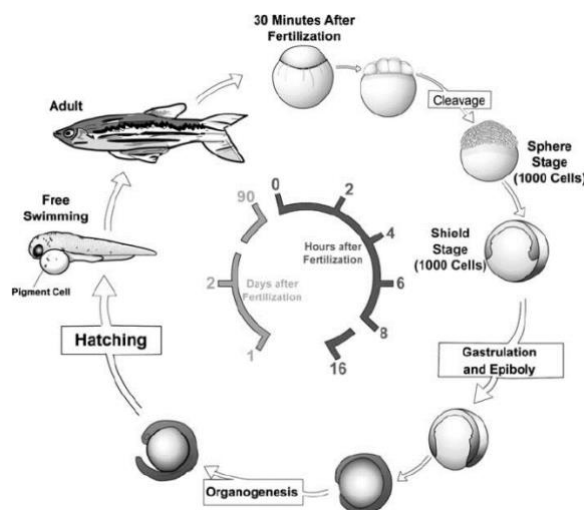


Figure 11. Cycle de vie du poisson zèbre (D'Costa and Shepherd, 2009).

2.5.2. Intérêt pour étudier la toxicité développementale

Les avantages expérimentaux

Le poisson zèbre est un modèle prisé en toxicologie et pharmacologie pour étudier les mécanismes d'action des composés et leurs effets sur le développement embryonnaire. Les raisons de cet attrait sont multiples puisqu'il a de nombreux avantages réglementaires et expérimentaux (Bambino and Chu, 2017; Larisch et al., 2017). Le poisson zèbre est une espèce

robuste facile à élever en laboratoire et peu couteuse comparé aux autres espèces animales (de Souza Anselmo et al., 2018a; Gonçalves et al., 2020). Sa fécondité est importante puisqu'un couple de poissons adultes est capable de pondre 200 à 300 œufs en une matinée, tous les 5 à 7 jours (Goldsmith and Jobin, 2012; Hill et al., 2005). Espèce ovipare, la fertilisation des œufs ainsi que leur développement se font hors de l'organisme maternel, facilitant l'accès à l'observation et la manipulation des embryons (Spitsbergen and Kent, 2003). Le développement est extrêmement court, permettant d'avoir un nombre conséquent d'individus pour les tests de toxicité. En 24 heures, les embryons ont un axe corporel défini et des organes rudimentaires, y compris un cœur contractile. De plus, l'âge adulte étant atteint à 3 mois de vie environ, la temps de génération de l'espèce est également rapide (Holtzman et al., 2016). La transparence des œufs et des larves permet la détection visuelle des malformations, sans interrompre le développement ni interventions invasives, atout majeur pour étudier notamment les effets tératogènes suite à l'exposition prolongée à un composé (Bernut et al., 2015; Martinez et al., 2018). Par exemple, des lignées transgéniques marquées avec un fluorophore sont maintenant disponibles pour une visualisation non-invasive des interactions pathogènes – macrophages/neutrophiles (Bernut et al., 2015; van der Vaart et al., 2012), de même que des gènes rapporteurs pour étudier l'expression de protéine suite à l'exposition d'un composé (Brion et al., 2012). La petite taille des embryons et des larves, 3,5 mm à 72 heures post fécondation, et leur capacité à vivre dans de faibles volumes d'eau permettent de réaliser des criblages haut-débit pour l'étude de la toxicité des composés, sur des plaques à puits (Fleming et al., 2013). Le génome du poisson zèbre, entièrement séquencé, peut être génétiquement manipulé pour générer des individus plus ou moins sensibles aux conditions testées. Il est possible, par blocage ou surexpression de gènes, d'étudier l'effet de mutations génétiques causées par les maladies ou l'exposition à un composé (Gonçalves et al., 2020; Howe et al., 2013).

Les homologies avec l'Homme

Au cours des 30 dernières années, le poisson zèbre est devenu un organisme modèle vertébré largement utilisé pour étudier la génétique, le biologie développementale, la neurophysiologie et la biomédecine, à l'instar des modèles mammifères classiques (Holtzman et al., 2016; Spence et al., 2007). Afin de justifier et de valoriser l'utilisation du poisson zèbre comme modèle prédictif de la toxicité humaine, il est primordial d'identifier les homologies mais également les différences entre les deux espèces (MacRae and Peterson, 2015).

Malgré l'existence de différences majeures résultant de l'adaptation à la vie aérienne, la plupart des organes du poisson zèbre remplissent les mêmes fonctions que leurs homologues humains et présentent une physiologie bien conservée, parfois mieux que chez les rongeurs. Au cours de la dernière décennie, de nombreuses pathologies digestives ont été modélisées chez le poisson zèbre car son système gastro-intestinal est doté d'un foie, du pancréas, de la vésicule biliaire et d'une voie intestinale segmentée linéairement avec des fonctions d'absorption et de sécrétion semblables à l'Homme (Goldsmith and Jobin, 2012; Zhao and Pack, 2017).

Le génome du poisson zèbre est composé de 25 paires de chromosomes contre 23 pour l'Homme (Shehwana and Konu, 2019). Le séquençage du génome du poisson zèbre a montré que 71% des protéines humaines et 82% des protéines humaines pathogènes ont un orthologue évident chez le poisson zèbre (MacRae and Peterson, 2015). Deux gènes sont dits orthologues s'il proviennent d'un ancêtre commun et qu'ils ne résultent pas d'une duplication génétique (Gabaldón and Koonin, 2013). Des protéines orthologues proviennent de gènes orthologues. Les protéines de poisson zèbre sont raisonnablement similaires à leurs orthologues humains, en particulier dans les domaines fonctionnels (Howe et al., 2013). Les modifications génétiques pouvant être effectuées facilement, de nombreux modèles de maladies humaines, utiles pour évaluer le potentiel thérapeutique et la toxicité des composés, ont pu être élaborés (Menke et al., 2011). De plus, le système immunitaire inné et adaptatif du poisson zèbre est similaire à celui de l'Homme. La fonctionnalité, l'activité et les profils d'expression de nombreux médiateurs cellulaires sont relativement conservés entre les deux espèces. Des études d'exposition hôte-pathogène mettant en jeu l'immunité innée peuvent être réalisées sur l'embryon puisqu'elle apparaît environ un jour post fécondation. L'immunité adaptative n'est pas analysable chez les embryons car elle n'apparaît qu'au bout de trois semaines de développement (van der Vaart et al., 2012). Le poisson zèbre est donc un modèle pertinent pour étudier les réponses immunitaires (Brugman, 2016; van der Vaart et al., 2012; Varela et al., 2012). Ses hépatocytes étant fonctionnels à partir de 36 heures post-fécondation, l'embryon est également une alternative aux modèles *in vivo* pour les tests d'hépatotoxicité (Driessen et al., 2013). Enfin, l'embryon de poisson zèbre est également utilisé comme modèle pour la toxicité développementale des vertébrés puisque la signalisation cellulaire responsable de l'embryogénèse est conservée (Martinez et al., 2018; Sipes et al., 2011a).

En évaluation des risques, déterminer la pharmacocinétique des composés est essentiel. Différentes études ont montré des homologies de partage des molécules entre les différents compartiments de l'organisme de poisson zèbre et de l'Homme. D'une part, passivement sur la

base des propriétés physico-chimiques des composés et d'autre part, activement par conservation de transporteurs actifs tissu-spécifiques au niveau des barrières physiologiques, telles que la barrière hémato-encéphalique (Farber, 2001; Fleming et al., 2013; Mihaljević et al., 2017; Popovic et al., 2014). De plus, bien que nous n'en soyons qu'au début de leur exploration, des processus métaboliques semblables ont été expérimentalement mis en évidence entre les deux espèces, incluant des cytochromes et d'autres enzymes. Des analyses génomiques ont établi la présence des gènes du cytochrome P450 chez le poisson zèbre (de Souza Anselmo et al., 2018a; Goldstone et al., 2010; Saad et al., 2017).

Malgré la mise en évidence de nombreuses similitudes, les auteurs s'accordent pour dire que le poisson zèbre est un modèle prometteur mais encore peu développé du fait de son intérêt récent. Beaucoup d'expérimentations visent à enrichir les connaissances biologiques pour accroître la puissance et la pertinence de l'utilisation du poisson zèbre pour modéliser la toxicité humaine, en complément ou remplacement des modèles mammifères.

2.5.3. Protocole d'étude de l'OCDE

Selon la Directive Européenne 2010/63/EU relative à la protection des animaux utilisés à des fins scientifiques (EU, 2010), les embryons de poisson zèbre ne sont pas considérés comme protégés ni soumis aux restrictions législatives jusqu'au cinquième jour post fécondation, marquant leur indépendance alimentaire de l'apport énergétique du sac vitellin (Braunbeck et al., 2015; Strähle et al., 2012). Par exemple, pour étudier la toxicité aigüe ou létale, les tests de toxicité sur les embryons de poisson (FET) sont proposés comme méthodes alternatives aux expérimentations animales. Deux applications scientifiques en tirent notamment avantages : l'identification de contaminants environnementaux ou de médicaments à potentialité tératogène pour l'Homme et la prédiction de la toxicité aigüe dans l'évaluation des risques environnementaux et humains (Klüver et al., 2014).

La Figure 12 décrit les stades de développement des poissons en relation avec le statut protégé et non protégé conformément à la Directive Européenne (EU, 2010), ainsi que les périodes d'application des tests de toxicité standard. Les FET standard incluent les embryons de 0 à 4 jours post fécondation. Cependant, pour le poisson zèbre, la durée peut être étendue jusqu'à 5 jours post-fécondation, considérant le stade éleuthéro-embryonnaire, i.e. entre l'éclosion et le stade larvaire (Braunbeck et al., 2015; Embry et al., 2010).

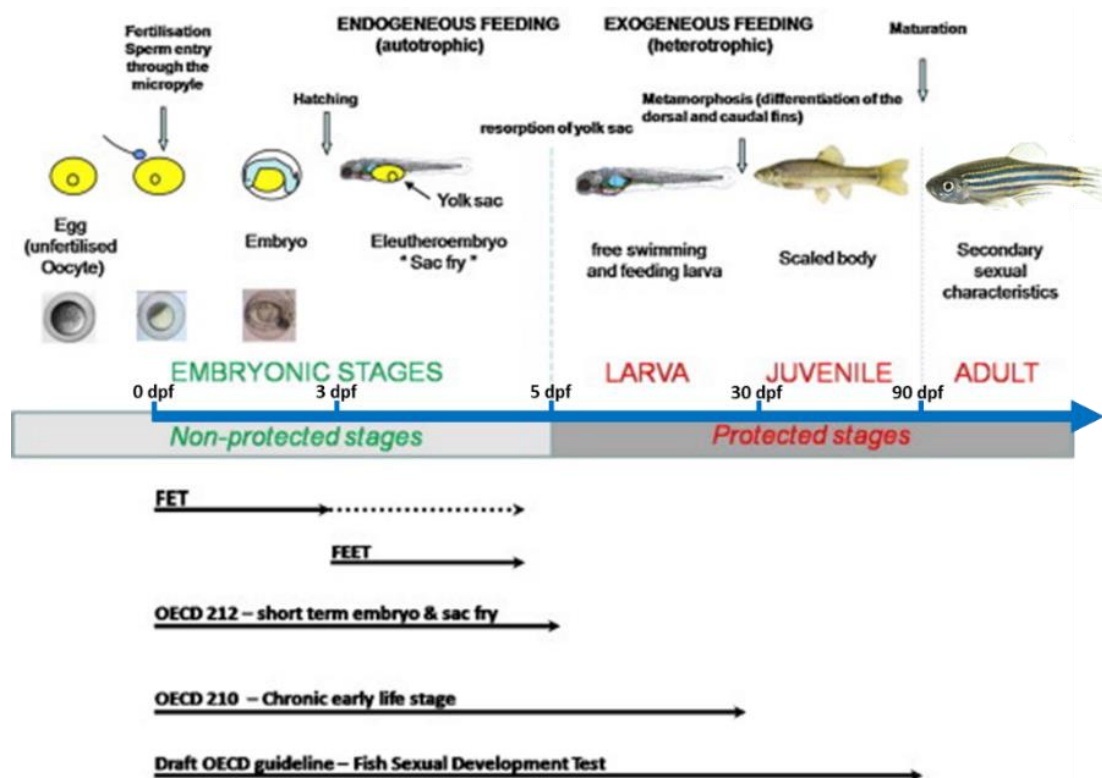


Figure 12. Stades de vie des poissons, statut de protection et durée des tests de toxicité standard. Les flèches pleines désignent la durée des tests référencés; la flèche en pointillés indique la durée supplémentaire autorisée pour les tests de l'embryon de poisson «étendu», comprenant le stade éleuthéro-embryonnaire. Adapté de (Embry et al., 2010).

2.5.4. Les limites des données des tests sur les embryons de poisson zèbre

L'extrapolation des résultats de toxicité sur l'embryon de poisson zèbre à l'Homme est soumise à plusieurs limites évidentes, incluant des facteurs cinétiques et dynamiques. La première limite majeure est la voie d'absorption, restreinte à la voie dermale, puisque les branchies n'apparaissent qu'au cinquième jour de développement. De plus, le chorion peut être considéré comme une barrière protectrice (mais ce n'est pas prouvé). Certains protocoles d'études le soustraient mais cela n'est pas sans conséquence sur le maintien de l'intégrité et du comportement de l'embryon, et prive également de la possibilité de déterminer les effets sur l'éclosion (Sipes et al., 2011a). Ensuite, le métabolisme d'un médicament peut varier considérablement selon le stade de développement de l'embryon de poisson zèbre (Alderton et al., 2010; Goldstone et al., 2010; Klüver et al., 2014). De plus, peu d'études sur la filtration, la réabsorption ou l'excrétion rénale des composés ont été entreprises à ce jour sur le poisson zèbre adulte et encore moins chez l'embryon. Enfin, malgré la disponibilité de techniques d'imagerie

performantes, la faible taille de l'embryon de poisson zèbre rend sa manipulation délicate, impactant la reproductibilité des études de dose-réponse puisque le comportement cellulaire et la viabilité peuvent être altérés. De plus, en raison des protocoles d'études imposant la fréquence d'analyse et du développement embryonnaire extrêmement court, certains effets se produisant pourraient ne pas être observés (Sipes et al., 2011a).

2.5.5. Modèles mathématiques pour le poisson zèbre

Les modèles mécanistes sont des outils utiles pour l'évaluation des risques des composés en complément des expérimentations (Hommen et al., 2016). En raison des nombreux avantages présentés précédemment à l'utilisation du poisson zèbre pour étudier la toxicité humaine, de nombreux modèles PK, PD, PBPK, et PK/PD ont été développés pour quantifier les processus toxiques déterminés expérimentalement et fournir les informations mécanistes nécessaires à l'extrapolation des données à l'Homme. Le Tableau 1 illustre le large domaine d'application des modèles permettant d'étudier la toxicité de différents composés.

Tableau 1. Tableau résumé des principaux domaines d'application des modèles PK, PBPK et PK/PD développés pour le poisson zèbre.

Modèle	Stade de vie	Domaine d'application	Caractéristiques	Auteurs
Pharmacocinétique compartimental	Larvaire et adulte	Substances psychoactives	Absorption et élimination : PK mono compartimentale Distribution : PK bi – compartimentale	(Kirla et al., 2018) (Kirla et al., 2016)
		Métabolisme	Absorption et distribution : PK mono compartimentale	(Van Wijk et al., 2019)
			Absorption et distribution : PK mono compartimentale	(Kantae et al., 2016)
	Embryon	Tératogénicité	Absorption et élimination : PK mono compartimentale	(Nawaji et al., 2018a)
		Pesticides	Absorption et élimination : PK bi – compartimentale	(Liu et al., 2018)
		Composés polaires	Absorption et élimination : PK mono compartimentale – facteur cinétique de bioconcentration	(Brox et al., 2016a)
Pharmacocinétique basé sur la physiologie	Larvaire et adulte	Pesticides	Circulation hépatobiliaire	(Khazaee and Ng, 2018)
		Perturbateurs endocriniens	Gonades	(Péry et al., 2014)
		Composés organiques neutres	Modèle multi – espèces	(Brinkmann et al., 2016)
		Toxiques environnementaux	Modèles multi – espèces Variabilité inter – espèces Facteurs environnementaux	(Grech et al., 2019)
	PBPK embryon	Tératogénicité	Clairance et croissance des organes	(Siméon et al., 2020)

Pharmacodynamique	Larvaire et adulte	Tératogénicité	Dose-réponse	(Gaballah et al., 2020)
	Embryon	Tératogénicité	Dose-réponse	(Brannen et al., 2010)
Pharmacocinétique/ Pharmacodynamique	Larvaire et adulte	Métaux	PK bi – compartimentale PD : modèle de survie	(Gao et al., 2019)
			PK mono compartimentale PD : modèle de survie	(Feng et al., 2018)
	Embryon	Tératogénicité	PBPK/PD multi – états	(Article chap. IV)

2.5.6. Synthèse des ZET

La pertinence des expérimentations animales ne permet pas de s'en contenter pour prédire la toxicité humaine mais elles représentent un atout non négligeable pour initier la connaissance du caractère toxique d'un composé, et son comportement au sein d'un organisme considéré dans son entier (van der Worp et al., 2010). Malgré les lacunes du poisson zèbre comme modèle pour prédire la toxicité humaine, il permet de faire le pont entre les études *in vitro* et les expérimentations animales, atout considérable notamment pour l'application de la réglementation des 3R. En effet, il peut être utilisé pour remplacer certaines études de toxicité animale, réduire le nombre d'animaux par identification en amont de composés toxiques à exclure d'emblée des expérimentations animales, et raffiner les conditions de vies des animaux en laboratoire, puisque la fécondation externe et la transparence des embryons de poissons zèbre permettent une observation non invasive (Cassar et al., 2020).

3. Les composés cibles en toxicologie développementale

Quotidiennement, les Hommes sont exposés à une variété croissante de substances toxiques naturelles, artificielles, chimiques ou biologiques. Les toxiques sont par définition des substances dangereuses, altérant le fonctionnement normal d'un organisme vivant par action

locale ou systémique (absorption cutanée, respiratoire, digestive). Il existe différentes mesures pour estimer la toxicité des composés : la dose ou concentration létale entraînant 50% de morts au sein d'une population (DL_{50} et CL_{50}), la dose minimale pour observer un effet toxique (LOAEL) et la dose sans effet toxique observable (NOAEL). La large utilisation des toxiques dans différents secteurs d'activités, tels que l'agriculture, la médecine, l'alimentation, l'industrie, *etc.*, prouve l'importance scientifique à déterminer, caractériser et quantifier la toxicité. En ce sens, la toxicologie prédictive trouve particulièrement son intérêt, notamment pour évaluer les propriétés et mécanismes d'action de substances inconnues. Ainsi, en évaluation de risques, une enquête approfondie sur les composés à potentialité toxique est primordiale pour élaborer une collection puis une classification des composés (Lapointe et al., 2004; Schmidt et al., 2009).

Les tests de toxicité pour le développement et la reproduction (DART) des produits pharmaceutiques, chimiques industriels et agrochimiques sont gouvernés par des directives harmonisées et mise en place à l'échelle mondiale dans le but de contrôler les risques d'exposition de différentes populations, notamment des femmes enceintes et enfants. Depuis leur adoption dans les années 1980 et 1990, ils font l'objet d'importants changements en raison des nouvelles préoccupations et considérations scientifiques. Les modifications sont notamment relatives à la volonté éthique de suivre la réglementation des 3R ainsi que l'addition de paramètres pertinents pour caractériser, déterminer et contrôler les perturbateurs endocriniens, avec pour ambition d'évaluer les dangers et les risques sans recours à l'expérimentation animale. Comme présenté dans les parties précédentes, les innovations scientifiques et techniques tels que les tests basés sur les cellules humaines, les modèles non mammifères, les tests à haut débit, les approches omiques, la biologie des systèmes et la modélisation informatique fournissent une base pour remplacer les tests sur les animaux. Les scientifiques de tous domaines investissent beaucoup d'efforts pour utiliser et optimiser les méthodes alternatives afin d'évaluer la toxicité des composés de manière générale, mais particulièrement la toxicité développementale spécifique au développement embryonnaire et fœtal, requérant un nombre important d'animaux pour mener à bien les études. En évaluation de risque développemental, deux catégories de composés, les perturbateurs endocriniens et les médicaments tératogènes, sont connus pour causer des malformations et altérer la croissance des embryons et des fœtus (Beekhuijzen, 2017).

Par exposition professionnelle, alimentaire et environnementale, les humains côtoient de nombreux produits chimiques identifiés comme perturbateurs endocriniens. Il s'agit de

composés altérant le fonctionnement normal du système endocrinien nécessaire au maintien de l'homéostasie, la reproduction, le développement et le comportement des animaux et humains. Les perturbateurs endocriniens sont hétérogènes et peuvent être classés de deux façons :

- Premièrement, selon leur catégorie, à savoir s'ils sont naturellement présents dans les aliments (ex : phytoestrogène) ou synthétisés (solvants ou lubrifiants industriels et leurs sous-produits (ex : dioxines), les plastiques (ex : Bisphénol A), les plastifiants, les pesticides (ex : dichlorodiphényltrichloroéthane (DDT)), les fongicides (ex : Vinclozoline) et certains agents pharmaceutiques (ex : diéthylstilbestrol)).

- Deuxièmement, selon leur origine, à savoir, les hormones naturelles et artificielles (ex : acides gras 3-oméga et pilules contraceptives), les médicaments à effets secondaires hormonaux (ex : métoprolol), les produits chimiques industriels et ménagers (ex : phtalates) et les produits secondaires des procédés industriels et ménagers (ex : hydrocarbures aromatiques polycycliques) (Kabir et al., 2015).

Depuis la tragique affaire du Thalidomide prescrit comme antiémétique aux femmes enceintes, puis reconnu comme responsable de malformations congénitales graves, l'identification du mécanisme d'action et du caractère tératogène des médicaments fait l'objet d'une attention particulière lors du développement de nouveaux composés mais également lors de l'utilisation de composés à structures chimiques semblables (Vargesson, 2015). Le défaut du tube neural (NTD) est considéré comme un des effets DART critiques observés *in vivo*. Dans cet optique, une étude de cas du projet européen *EuToxRisk*, intégrant ces travaux de thèse, s'est spécifiquement intéressée à l'acide valproïque (VPA), acide gras à chaîne courte, provoquant des NTD chez les animaux et les humains, dont la formule chimique est présentée par la Figure 13. Des études chez l'animal ont montré que certains des neuf analogues du VPA induisaient également des NTD (acide 2,2-diméthylvalérique, acide 2-éthylbutyrique, acide 2-éthylhexanoïque, acide 2-méthylhexanoïque, acide 2-méthylpentanoïque, acide 2-propylheptanoïque, 4-eneVPA, acide 4-pentenoïque and acide hexanoïque).

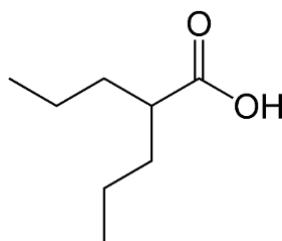


Figure 13. Structure chimique de l'acide valproïque.

Le VPA est largement prescrit par administration orale et intraveineuse pour contrôler les crises d'épilepsie et réguler la phase maniaque associée au trouble bipolaire. Son activité anticonvulsivante est régie par l'inhibition du cycle de l'acide citrique et l'élévation du taux d'acide γ -aminobutyrique (GABA) au sein du système nerveux central. Le taux sanguin thérapeutique de VPA est généralement compris entre 0,3 et 0,6 mM, mais les ajustements dépendent de la réponse clinique (Sztajnkrycer, 2002). Pour cette posologie, la concentration cérébrale en VPA est d'environ 0,5 mM signifiant que la molécule est largement transportée à travers la barrière hémato-encéphalique. La Figure 14 montre plus précisément le mécanisme d'action du VPA. L'acide α -cétoglutarique est généré à partir du cycle de l'acide tricarboxylique (TCA) et peut être transformé en métabolites selon deux voies différentes. La première le métabolise en succinyl-coenzyme A par action de l' α -cétoglutarate déshydrogénase puis en succinate par la succinyl-coenzyme A synthétase. La deuxième le métabolise successivement en glutamate par la glutamate oxaloacétate transaminase, en GABA par la glutamate décarboxylase, en succinate semi-aldéhyde par la GABA transaminase puis en succinate par la semi-aldéhyde déshydrogénase (Chuang et al., 2012).

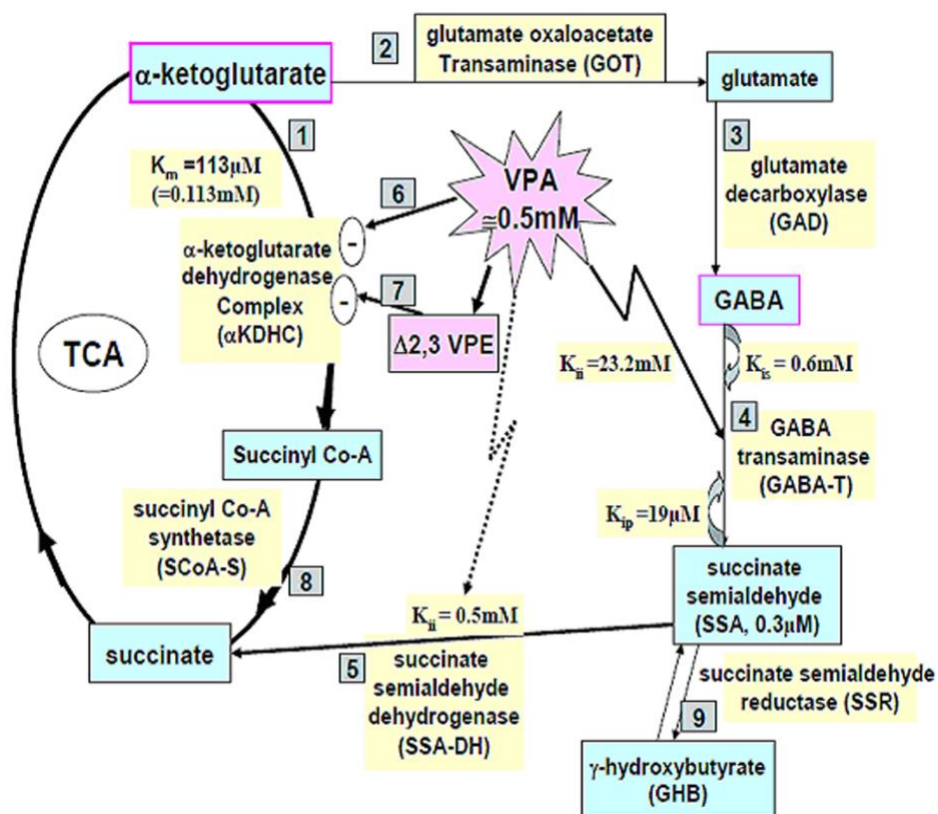


Figure 14. Mécanisme d'action pharmacologique de l'acide valproïque (Chuang et al., 2012).

Fin 2009, l'Agence américaine des produits alimentaires et médicamenteux (FDA) publie un avertissement indiquant le risque important de malformations congénitales majeures, notamment des anomalies crânio-faciales, neurales, cardiaques et cognitives, associées à la prise de VPA pendant la grossesse (Fathe et al., 2014). Des recherches récentes ont permis de confirmer la tératogénicité du composé mais également le risque de perturbation hormonale. L'exposition *in utero* augmente le risque d'une variété de malformations congénitales majeures, d'un quotient intellectuel réduit et de problèmes comportementaux. Chez les filles et femmes, le VPA augmente le risque d'anomalies hormonales, d'obésité et de syndrome des ovaires polykystiques (Gotlib et al., 2017). Le mécanisme exact responsable de sa tératogénicité n'est pas clarifié et pourrait être le résultat de différentes actions conjointes ou isolées. Les hypothèses actuelles s'orientent principalement vers l'inhibition de l'histone désacétylase, l'interférence avec le métabolisme des folates et l'induction du stress oxydatif (Chuang et al., 2012). Malgré les directives de mise en garde et réglementations actuelles, le VPA continue d'être prescrit aux femmes en âge de procréer du fait de son efficacité (Bosak et al., 2019; Gotlib et al., 2017). L'avis de l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (Anses) relatif à une demande d'analyse critique des valeurs toxicologiques de référence du VPA a établi qu'au vu des connaissances actuelles, aucune NOAEL ne peut être définie. Sans certitude, la dose thérapeutique initiale, 10 mg/kg/j, est la LOAEL de référence pour la femme enceinte, sachant qu'elle ne permet pas d'exclure le risque tératogène (ANSES, 2018).

En conclusion, nous avons dans un premier temps présenté dans l'état de l'art les principaux modèles pharmacocinétiques et pharmacodynamiques utilisés en évaluation des risques pour extrapoler les données de toxicité *in vitro* et animales à l'Homme. Ensuite, nous avons énoncé les différents types de données de toxicologie développementale disponibles pour l'élaboration des modèles mathématiques, et particulièrement les données obtenues suite à des tests de toxicité réalisés sur les embryons de poisson zèbre. Pour finir, nous avons succinctement présenté les composés cibles actuellement en toxicologie développementale, notamment l'acide valproïque, composé étudié pour les travaux de cette thèse. Dans la suite, seront présentés les modèles de distribution virtuelle *in vitro*, PBPK et PK-PD multi-états développés pour prédire la toxicité développementale de l'acide valproïque chez l'embryon de poisson zèbre.

CHAPITRE II. Modèle de distribution *in vitro* VIVD

1. Un nouveau modèle de distribution virtuelle *in vitro*

1.1. Objectif des travaux et contribution

Dans chaque test *in vitro*, plusieurs facteurs peuvent influencer le comportement des composés étudiés, puisque selon ses propriétés physico-chimiques, une molécule peut par exemple se fixer au plastique des contenants, aux composants du milieu de culture, s'accumuler plus ou moins dans les cellules test ou s'évaporer. Tous ces facteurs peuvent impacter de façon non-linéaire la concentration de composé dans les cellules test. Cette concentration, qui peut varier dans le temps, conditionne les effets observés. Il est donc important, pour comparer les résultats de différents tests *in vitro*, ou bien les transposer à l'animal ou l'humain, de connaître la concentration cellulaire du composé testé et d'établir une relation entre cette concentration et les effets observés. Il n'est pas toujours facile de mesurer la concentration cellulaire de composé, et souvent, seule la concentration nominale dans le système *in vitro* (quantité de produit introduite sur volume total du système) est connue. Même si les concentrations cellulaires ont été mesurées, leur analyse statistique devrait être étayée par un modèle de la relation entre concentration nominale et concentration cellulaire. C'est précisément pour prédire les concentrations intracellulaires de produits chimiques dans les essais *in vitro* que le modèle de distribution virtuelle *in vitro* a été développé par Ciaran Fisher de Certara, partenaire du projet EU-ToxRisk (Fisher et al., 2018). Il est applicable à la plupart des petites molécules ionisables ou non.

Ces travaux de recherche ont été valorisés dans un article « *VIVD: virtual in vitro distribution model for the mechanistic prediction of intracellular concentrations of chemicals in in vitro toxicity assays* », publié dans *Toxicology in Vitro* en 2018 (DOI: 10.1016/j.tiv.2018.12.017) et présenté ci-après. Le matériel supplémentaire est disponible en Annexe 1.

Notre participation à ces travaux a consisté à rendre dynamique le modèle initialement décrit à l'état d'équilibre, en tenant compte de la quantité de composé progressivement métabolisée (et donc éliminée) au cours du temps. Les équations statiques des coefficients de partage ont donc été couplées à une simple équation différentielle décrivant le changement de

la quantité totale de produit dans le système. Ce modèle était un prérequis à la construction et au fonctionnement de notre modèle PBPK (dynamique) de l'embryon de poisson zèbre, présenté en chapitre 3.

1.2. Perspectives

L'approche présentée ici est relativement pragmatique comparée à des modèles plus complexes, puisque la distribution *in vitro* prédite à l'état d'équilibre peut introduire un biais selon la nature du système cellulaire utilisé. Seule une vérification limitée de ce modèle a été publiée à ce jour et des travaux expérimentaux supplémentaires sont en cours au sein du projet EU-ToxRisk. Cependant, ce modèle nous a permis de déterminer les coefficients de partage de l'acide valproïque et neufs de ses analogues selon leurs propriétés physico-chimiques et les caractéristiques physiologiques de chaque organe de l'embryon de poisson zèbre étudiés. Ils constituent les paramètres d'entrée du modèle PBPK embryon de poisson zèbre présenté en détails dans le chapitre suivant (chapitre 3).

2. Résumé et article du modèle VIVD

Les tests de toxicité *in vitro* utilisent généralement des concentrations nominales de traitement pour déterminer les paramètres de toxicité. Cependant, les composés étudiés peuvent se lier au plastique des récipients de culture cellulaire ou interagir avec les composants du milieu de culture, tels que les lipides et l'albumine. De plus, les composés volatils peuvent s'évaporer dans l'air au-dessus du milieu de culture. Ces processus réduisent les concentrations libres de composé auxquelles les cellules sont exposées.

Des modèles prédictifs des concentrations libres dissoutes et considérant les différentes interactions des composés avec les composants de l'environnement d'étude *in vitro* ont été publiés. Cependant, ceux-ci n'ont été appliqués qu'aux composés neutres ou ne supposent aucune ionisation différentielle entre le milieu et le cytoplasme cellulaire. Les travaux de cet article décrivent un modèle de distribution *in vitro*, basé sur l'équation de Fick-Nernst Planck tenant compte de l'ionisation différentielle des composés dans le milieu de culture et l'eau intracellulaire. Le modèle considère la perméabilité des espèces ionisées et non ionisées, ainsi que le potentiel membranaire impactant le partage des groupements ionisés. En tenant compte de la liaison lipidique et protéique dans le milieu de culture, la liaison au plastique du système de culture cellulaire, la répartition de l'air et la liaison lipidique dans la cellule, le modèle peut

prédire les concentrations chimiques (libres et totales) dans le milieu de culture et les cellules. Le modèle peut améliorer l'extrapolation *in vitro* – *in vivo* des paramètres de toxicité en déterminant les concentrations intracellulaires pour la transposition *in vivo*.

VIVD: virtual *in vitro* distribution model for the mechanistic prediction of intracellular concentrations of chemicals in *in vitro* toxicity assays*

Fisher C^{a*}, Siméon S^b, Jamei M^a, Gardner I^a, Bois Y. F^b

^a Certara UK Limited, Simcyp Division, Acero, 1 Concourse Way, Sheffield S1 2BJ, UK; Email: *ciaran.fisher@certara.com

^b INERIS, METO Unit, Verneuil en Halatte, France

Abstract

In vitro toxicity testing routinely uses nominal treatment concentrations as the driver for measured toxicity endpoints. However, test compounds can bind to the plastic of culture vessels or interact with culture media components, such as lipids and albumin. Additionally, volatile compounds may partition into the air above culture media. These processes reduce the free concentrations of compound to which cells are exposed.

Models predicting the freely dissolved concentrations by accounting for these interactions have been published. However, these have only been applied to neutral compounds or assume no differential ionisation of test compounds between the media and cell cytoplasm. Herein, we describe an *in vitro* distribution model, based on the Fick-Nernst Planck equation accounting for differential compound ionisation in culture medium and intracellular water. The model considers permeability of ionised and unionised species and accounts for membrane potential in the partitioning of ionised moieties. By accounting for lipid and protein binding in culture medium, binding to cell culture plastic, air-partitioning, and lipid binding in the cell, the model can predict chemical concentrations (free and total) in medium and cells. The model can improve *in vitro in vivo* extrapolation of toxicity endpoint by determining intracellular concentrations for translation to *in vivo*.

Keywords

in vitro assays, IVIVE, biokinetics, toxicity

* Toxicology in Vitro. 58 (2019) 42–50. <https://doi.org/10.1016/j.tiv.2018.12.017>.

1. Introduction

Integrated testing strategies (ITSs), combining *in vitro* assay systems, *in silico* methodologies, knowledge-bases, and omics driven mechanistic insights are increasingly replacing the use of animal models in toxicological risk assessment. This has been driven, not only by the increasing availability and applicability of these technologies, but also by an increasing societal and legislative impetus to refine, reduce and, ultimately, replace animal testing. The 3Rs principle, applied to the use of animal models, is not solely based on ethical considerations of animal welfare, but also on the increasing acceptance that extrapolation from preclinical species to humans is not sufficiently robust for risk assessment in the 21st century (Gibb, 2008; Holmes et al., 2010).

Cell based *in vitro* systems can be used to determine the concentration dependant response for a range of toxicological endpoints, from simple cytotoxicity assays, to highly tissue-specific functional or transcriptomic responses (Godoy et al., 2013; Meek and Lipscomb, 2015). The use of such assays seeks to inform a more mechanistic understanding of the pathways that mediate adverse outcomes (Leist et al., 2017). *In vitro* assays are ideally placed to operate in this way since target-organ, species-specific material can be used in these model systems, and they are amenable to scaling to high-throughput screening approaches. The Tox21 project seeks to utilise *in vitro* assays in such an approach in order to assess the hazard for a huge number of compounds for which *in vivo* data is unavailable and realistically unobtainable (Tice et al., 2013). Based on the data generated, an *in vitro in vivo* extrapolation (IVIVE) approach can then be adopted in assessing human and animal hazard (Bell et al., 2018).

However, toxicological endpoints quantified *in vitro* are routinely correlated to a range of nominal treatment concentrations to establish a dose-response relationship. Yet, using the nominal treatment concentration as the driving concentration for toxicity is inaccurate since it does not account for the differential distribution of compounds within the test system that will determine the true free concentration available for distribution into the cell, and subcellular compartments, and so determine the driving concentration for toxicity at the target site. In order to have a more accurate assessment of concentration driven toxicity, it is necessary to correct the nominal treatment concentration, accounting for the different factors affecting a compounds distribution. These include binding to the plastics used in cell culture, exchange at the interface between culture medium and the air in the headspace, and binding to components within serum

that may be included in culture medium (e.g. lipids and proteins); the modelling of these processes is termed biokinetics (Blaauboer, 2010).

A number of steady-state biokinetic models have been published that account for some or all of the factors listed above. Armitage and colleagues (Armitage et al., 2014) published a model framework to predict intracellular concentrations, correcting for some of the distribution factors in monolayer cell culture. This steady-state framework assumes instantaneous partitioning between media, headspace, serum-lipids, serum-proteins, dissolved organic material, and the cultured cell volume. The steady-state assumption lends itself to the modelling of non-proliferating cells, at the time of the assay, with minimal metabolic activity against the test compound. Moreover, a critical assumption of the Armitage model is that the test compounds are neutral or not significantly ionised under the conditions of the *in vitro* assay. This assumption of neutrality was, to some degree, addressed in the model developed by Fischer *et al* (Fischer et al., 2017) where the authors adopted the same steady-state assumption, but excluded the partitioning of compound into the headspace. The authors incorporated separate partition constants for both the ionised and unionised fraction of test compound, determining the fraction ionised assuming a uniform pH=7.4 throughout the test system. This neglects the differential ionisation potential between the culture media and intracellular water resulting from their differing pH. Furthermore, the interior of the cell itself is not a uniform environment with the interior of specific organelles being maintained at a specific pH critical to their function (i.e. lysosomes (pH≈4.5), mitochondria (pH≈8), cytosol (pH≈7)). Indeed, the differential ionisation of compounds between organelles and intracellular water can result in the preferential sequestration of compounds within organelles; a phenomena commonly known by the misnomer ‘ion-trapping’ (Kazmi et al., 2013). It is also critical to note that differences in the intrinsic permeability of the unionised/ionised form are not the only factors determining the distribution of ionised compound into cells. The potential difference maintained across the cell membrane, membrane potential (Φ , mV), can actively promote the uptake or exclusion of ionised compounds from the cell interior and can vary significantly between cell types. It is important to remember that these models (Armitage et al., 2014; Fischer et al., 2017) are steady-state approximations of multiple dynamic processes. Critically, they assume that the loss of test compounds within the *in vitro* system is negligible with no loss through metabolic clearance or instability. When that is not the case, the steady-state assumption may lead to overestimation of intracellular concentrations of the parent molecule, particularly for highly metabolized chemicals. Obviously, this restricts the applicability of these models and limits their use in the

assessment of ionisable or metabolically cleared compounds. Dynamic (ODE-based) models that move beyond the steady-state assumptions described have been published incorporating metabolic clearance and distribution into mitochondria (Worth et al., 2017; Zaldivar Comenges et al., 2017). However, these models still assume no significant ionisation of the test compound.

Here we propose two versions of an improved virtual *in vitro* intracellular distribution (VIVD) model: the first is a steady-state formulation which relates a molecule's distribution to its ionization state, membrane permeability, and volatility. The VIVD model also accounts for cell membrane potential, pH, and extracellular binding to culture medium proteins, lipids, and plastics, but metabolism is considered negligible. The second is a quasi-steady-state dynamic formulation incorporating metabolic clearance which can be used when metabolic loss of compound cannot be considered negligible. The impact of the various improvements proposed are illustrated through simulations.

2. Materials and Methods

The model structure described below assumes monolayer cell culture in wells (i.e. plate format) or in round dishes; this distinction is only important in considering test compound binding to plastic culture-ware and will be discussed below. The parameterization used herein assumes a monolayer culture of primary human hepatocytes. But the model can easily be adapted to predict distribution into other cell types by including cell specific system parameters (see below for further explanation).

The (*in vitro*) systems parameters for the model described herein, which are independent of the test compounds, can be subdivided into two categories; those related to cultured cells and those pertaining to the culture conditions. Each of these will be described below. Default units for calculations within the model are litres (L) for volume, metres (m) for length/area and molar for concentration (M).

2.1. Cell Composition

Hepatocytes account for ~60% of the cells in the liver and account for 80% of the parenchymal volume (Godoy et al., 2013). Here, we assume that viable isolated human hepatocytes will maintain a comparable size and composition with respect to lipids as to hepatocytes within human liver. The diameter of human hepatocytes has been cited as being between 20-40 μ m, while mouse hepatocytes have been reported as being 23.3 $\pm$ 3.1 μ m in diameter (Hoehme et al., 2010; Schwen et al., 2015). Here we take 30 μ m to be the diameter of

human hepatocytes in 2D culture, $Cell_{diam.}$ (μm). Assuming that cells cultured in 2D-monolayers take a dome-like shape (half-spheres) on the culture surface, and knowing the number of cells at the time of the assay, we can calculate the total volume of cells, $V_{total\ cell}$, using Eq. 1.

$$V_{total\ cell}(L) = n_{cell} \left(\frac{\frac{4}{3}\pi(Cell_{diam.} \cdot 0.5e^{-6})^3}{2} \right) \cdot 10^3 \quad (1)$$

2.2. Culture Conditions

The formats available for cell culture are diverse. However, the adoption of high-throughput and automated methods has necessitated a degree of standardisation across these formats. In most cases high-throughput assays are conducted in plate formats comprising cylindrical wells. The model assumes a hermetically sealed, cylindrical culture system corresponding to an individual well of a culture plate or a single culture dish. Based on this assumption, the model can be easily parameterised to reflect the diverse range of culture plate formats that are commercially available. Four culture condition parameters are required to define the system that is being used: diameter of a single culture vessel ($System_{diam.}$; mm), volume of the culture system (V_{well} ; μL), the volume of medium (V_{medium} ; μL) and the fraction of medium (v/v) comprising fetal bovine serum (FBS) (f_{serum}). More specialised culture mediums, where serum is pretreated (e.g. delipidized), or serum free systems where specific concentrations of lipid or albumin are added can also be simulated.

The volume of air (V_{air} ; L), above the medium can be calculated using Eq. 2 and is subsequently used to calculate the concentration of test compound partitioned into the headspace at steady state.

$$V_{air} = 10^{-6}(V_{well} - V_{medium}) - V_{totalcell} \quad (2)$$

To calculate the binding of test compound to plastic it is necessary to calculate the surface area of plastic in direct contact with the medium, $SA_{medium\ plastic}$ (m^2), Eq. 3. Here we assume a round well culture system and so the only variables will be the diameter of the well and the volume of medium.

$$SA_{medium\ plastic} = 2\pi r \left(\frac{V_{medium} \cdot 10^{-9}}{\pi(system_{diam.} \cdot 0.5e^{-3})^2} \right) \quad (3)$$

We assume that the cell monolayer cultured on the bottom of the well prevents any contact of the media with the plastic of the well bottom. This assumes 100% cell confluence; in dividing cell-lines this can result in contact inhibition of growth.

2.3. Binding to Serum

FBS is routinely included in cell culture media to supply growth factors, hormones and other factors that support cell growth and improve viability (Gstraunthaler, 2003). Specialist FBS preparations are also available that modify the composition of FBS. For example, charcoal stripped FBS containing no lipid. The major protein in FBS is albumin (>60% total protein) and this is routinely measured and detailed in the certificate of analysis (CoA) (Zheng et al., 2006). FBS also contains neutral lipids and free fatty acids unless specially treated; the plethora of lipids within FBS is not routinely quantified, but triacylglyceride (TAG) content is reported in the CoA. It is necessary to account for the binding of test compounds to these major serum components in order to determine the true-free concentration in the *in vitro* system. Where the fraction unbound in serum, f_{uFBS} is determined experimentally, it can be used directly in the model. Alternatively, f_{uFBS} can be predicted using Eq. 4.

$$f_{uFBS} = \frac{1}{1 + K_{protein}f_{protein} + \frac{P_{nl}f_{nl,FBS}}{X_{FBS}}} \quad (4)$$

Where $f_{protein}$ (v/v) is the fraction of FBS comprised of protein, $K_{protein}$ is the albumin:water partition coefficient, P_{nl} is the neutral lipid partition coefficient (defined below), $f_{nl,FBS}$ is the fraction of FBS comprised of neutral lipid and X_{FBS} is an ionisation term (Table 1), assuming that binding to neutral lipids is limited to unionised species.

Table 1. Calculation of Y and X ionisation terms using Henderson Hasselbalch (Berezhkovskiy, 2011)

Compound Type	Y	X
Neutral	$Y_1 = Y_2 = 0$	$X = 1$
Monoprotic acid	$Y_1 = 10^{pH-pKa}, Y_2 = 0$	$X = 1 + Y_1 + Y_2$
Monoprotic base	$Y_1 = 10^{pKa-pH}, Y_2 = 0$	$X = 1 + Y_1 + Y_2$
Diprotic acid	$Y_1 = 10^{pH-pKa1} + 10^{pH-pKa2},$ $10^{2pH-(pKa1+pKa2)}$	$Y_2 =$ $X = 1 + Y_1 + Y_2$
Diprotic base	$Y_1 = 10^{pKa1-pH} + 10^{pKa2-pH},$ $10^{(pKa1+pKa2)-2pH}$	$Y_2 =$ $X = 1 + Y_1 + Y_2$
Ampholyte	$Y_{1a} = 10^{pH-pKa1},$ $10^{pKa2-pKa1}$	$Y_{1b} = 10^{pKa2-pH},$ $Y_2 =$ $X = 1 + Y_{1a} + Y_{1b} + Y_2$

Ionisation terms are calculated for the FBS, culture medium, intracellular water, mitochondria and lysosome; for ampholytes, pKa1 corresponds to acidic group and pKa2 to the basic group.

Assuming that the fraction of albumin in the FBS is representative of the total protein fraction responsible for protein binding in the FBS, this can be calculated from the mass of albumin reported in the CoA for a batch of FBS using Eq. 5.

$$f_{protein} \approx f_{alb,FBS} = \frac{mass\ albumin \cdot PSV_{albumin}}{1000} \quad (5)$$

Where $PSV_{albumin}$ is the partial specific volume of albumin (0.73 mL/g (Kupke et al., 1972)); it should be noted that Eq. 5 does not in fact yield a true fraction, but the volume of albumin in litres. Here we assume one litre of FBS for the prediction of f_{uFBS} , and so calculation of the volume specific serum components in litres is equivalent to (v/v) fractions. The albumin to water partition coefficient, $K_{protein}$, can be determined experimentally or can be calculated as previously described (Endo and Goss, 2011) in Eq. 6; in this work the calculation approach has been used.

if $\log P_{ow} < 4.5$

$$\log K_{protein} = 1.08 \cdot \log P_{ow,Tsys} - 0.7$$

if $\log P_{ow} \geq 4.5$

$$\log K_{protein} = 0.37 \cdot \log P_{ow,Tsys} + 2.56 \quad (6)$$

In much the same way, Eq. 7 can be used to calculate a representative (v/v) fraction of neutral lipid in FBS, using TAG as a surrogate for neutral lipid content. TAG concentration is routinely determined using an enzymatic assay and so reported as a molar concentration.

$$f_{nl,FBS} \approx f_{TAG} = \frac{[TAG] \cdot 10^{-3} \cdot MW_{TAG} \cdot PSV_{TAG}}{1000} \quad (7)$$

Where PSV_{TAG} is the partial specific volume of TAG (1.09 mL/g (Deckelbaum et al., 1984)) and the molecular weight of TAG is taken to be 885.453 g/mol; specifically, this corresponds to the molecular weight of trioleate, a TAG molecule comprising a glycerol backbone and three oleic acid residues.

Irrespective of whether f_{uFBS} is determined experimentally or is predicted, its value represents the binding in serum alone. However, serum routinely only comprises a fraction of

the complete medium. Consequently, it is necessary to correct the predicted or experimentally determined value to account for the dilution of serum in total culture medium. Firstly, a dilution factor is calculated based on the fraction of serum, f_{serum} , in complete medium (Eq. 8) and then an f_{uFBS} value corrected for dilution, $f_{uFBS,dilu}$, calculated (Eq. 9).

$$D = \frac{1}{f_{serum}} \quad (8)$$

$$f_{uFBS,dilu} = \frac{f_{uFBS}}{\frac{1}{D}(1-f_{uFBS})+f_{uFBS}} \quad (9)$$

If $f_{serum} = 0$, then $f_{uFBS,dilu} = 1$.

2.4. Partitioning to Plastics

In addition to binding to serum components, test compounds may also bind to the plastic of the culture vessel. Here we employ a linear correlation between the plastic to media partition coefficient, $K_{plastic}$, and the octanol to water partition coefficient, P_{ow} , established by (Kramer, 2010) and later used by (Zaldivar Comenges et al., 2017). It *should* be noted, that this relationship was established for of polycyclic aromatic hydrocarbons (PAHs) and so may not hold true for all classes of compound. In the approach outlined here it is assumed that both unionised and ionised fractions of solubilised compound can bind to culture plastics (Eq. 10). It should also be noted that binding to serum components can be considered to be of greater significance than binding to plastics. Indeed, it has been demonstrated that the significance of binding to plastics is reduced with increasing supplementation of FBS in culture media (Kramer et al., 2012).

$$K_{plastic,u} = \frac{C_{plastic}}{C_{medium_{unbound}}} = 10^{0.97 \log P_{ow} T_{sys} - 6.94} \quad (10)$$

It is important to note that $k_{plastic}$, unlike other partition coefficients used herein, is not dimensionless and has units of metres (m).

2.5. Partitioning into Headspace Air

Cells are routinely cultured in humidified incubators to help limit the evaporation of media. The plate formats utilised in high-throughput assays do not routinely seal individual wells on the plate; lids either cover the entire plate, limiting evaporation, or are handled uncovered due to the use of automated fluidic systems. This means that the volume of air into which compounds solubilised in the media may distribute is poorly defined. The model described here does not take into account the mixing of air above a whole culture plate, or the

possible cross-contamination between wells that may occur if multiple volatile compounds or test concentrations are being tested across a single unsealed plate. Here we assume that each well of a plate format exists as an independent, hermetically sealed system.

The exchange of compound between the atmosphere and water is already an established concept in environmental toxicology (Schwarzenbach et al., 2004). At dilute concentrations in pure water, the distribution of neutral compounds between the air and the water is described by the Henry's law constant, K_H (Eq. 11).

$$K_H = \frac{P_i}{C_{water,unbound}} \quad (11)$$

Where P_i is the partial pressure of the gas phase of the compound (atmLmol^{-1}) and $C_{water,unbound}$ is the molar concentration in water. Here, it is more convenient to redefine the Henry's law constant of compounds as a dimensionless partition coefficient, $K_{air,u}$ (Eq. 12; also defined as K_{aw}).

$$K = \frac{K_H}{RT} \quad (12)$$

Where R is the universal gas constant and T is the reference temperature in Kelvin (normally $25^\circ\text{C} = 298.15\text{K}$). It is important to note that the terms, and therefore the units, in which Henry's law constants are reported are inconsistent and can lead to confusion. In Eq. 12 it is essential that the universal gas constant is expressed in units consistent with those of the Henry's law constant (i.e. if K_H is expressed in SI units of $\text{Pa m}^3 \text{mol}^{-1}$ then $R = 8.314 \text{ Pa m}^3 \text{K}^{-1} \text{mol}^{-1}$; if K_H is expressed in units of atm L mol^{-1} then $R = 0.08206 \text{ L atm mol}^{-1}$).

Besides databases listing experimentally determined Henry's law constant values, a number of tools are available for predicting this parameter (Dearden and Schuurmann, 2003; Katritzky et al., 1998). The HenryWin tool gives users a prediction of K_H in SI units, but also as the dimensionless $K_{air} = K_{air,u}$ partition coefficient (Dearden and Schuurmann, 2003).

As already stated above, the Henry's law constant describes the air-water distribution ratio of neutral compounds in non-saturated and pure water, single solute systems. As such, its use here is an approximation of the behaviour in a 'real' aqueous solution containing multiple chemical species. However, to limit the assumptions being made, $K_{air,u}$ is only applied to the unionised fraction, f_{ui} , of solubilised compound.

$$\log K_{air,Tsys} = \log K_{air} - \frac{\Delta U_{aw}}{\ln 10 \cdot R} \cdot \left(\frac{1}{T_{sys}} - \frac{1}{T_{ref}} \right) \quad (13)$$

Experimental, and predicted values, of Henry's law constant are reported at a reference temperature (T_{ref}) of 25°C, while the temperature cell culture systems (T_{sys}) is routinely maintained at 37°C. It is therefore necessary to correct the values used in the model to the relevant system temperature using Eq. 13, where ΔU_{aw} (J mol⁻¹) is the internal energy of phase change for air to water partitioning.

2.6. Partitioning in Cells

Using an adapted form of the Rodgers and Rowland approach (Rodgers et al., 2005; Rodgers and Rowland, 2007) we describe the partitioning of test compound into cells. This approach adopts a more mechanistic description of intracellular lipid binding and is based on the pH-partition hypothesis. It assumes that only the unionised fraction can partition across tissue membranes. Therefore, at steady-state, the concentration of unbound, unionised compound between the intracellular and extracellular water is equal (Eq. 14).

$$K_{cell\ uu,uu} = \frac{c_{intracellular}^{unbound,unionised}}{c_{extracellular}^{unbound,unionised}} = 1 \quad (14)$$

The unionised concentrations are defined by the Henderson-Hasselbalch equation based on compound type and the local pH (Table 1). In predicting the partitioning of neutral or weak acids/bases, the pH-partition is valid since the compound is unionised or the fraction of ionised compound is negligible and so makes no significant contribution to partitioning across membranes. However, as noted above the negative potential difference maintained across most cell membranes promotes the partitioning of cations into the cell and limits the partitioning of anions.

Using the Fick-Nernst-Planck equation, models for passive cell permeability accounting for differential ionization (Henderson-Hasselbalch equation), passive permeation of neutral molecules (Fick's law), and passive permeation of ionized molecules (Nernst-Planck equation) across the cell membranes have been published previously (Ghosh et al., 2014a; Ghosh et al., 2014b; Trapp et al., 2008).

Here we adapt the existing Rodgers and Rowland approach incorporating the Fick-Nernst-Planck equation to describe the passive penetration of electrolytes across cell membrane, together with the passive penetration of neutral molecules. $K_{cell\ uu,uu}$ is introduced and derived from the steady-state Fick-Nernst-Planck equation (Eq. 15, Eq. 16, Table 2) and further used to predict the cell to media partition coefficient (K_{cell} , Eq. 17).

$$K_{cell\ uu,uu} = \frac{1 + \frac{p_{unbound,ionised\ 1}}{p_{unbound,unionised}} \frac{N_1}{e^{N_1-1}} Y_{1,EW} + \frac{p_{unbound,ionised\ 2}}{p_{unbound,unionised}} \frac{N_2}{e^{N_2-1}} Y_{2,EW}}{1 + \frac{p_{unbound,ionised\ 1}}{p_{unbound,unionised}} \frac{N_1}{e^{N_1-1}} Y_{1,IW,1} + \frac{p_{unbound,ionised\ 2}}{p_{unbound,unionised}} \frac{N_2}{e^{N_2-1}} Y_{1W,2}} \quad (15)$$

$$K_{IW_{organelle}^{uu,uu}} = \frac{1 + \frac{p_{unbound,ionised\ 1}}{p_{unbound,unionised}} \frac{N_1}{e^{N_1-1}} Y_{1,IW} + \frac{p_{unbound,ionised\ 2}}{p_{unbound,unionised}} \frac{N_2}{e^{N_2-1}} Y_{2,IW}}{1 + \frac{p_{unbound,ionised\ 1}}{p_{unbound,unionised}} \frac{N_1}{e^{N_1-1}} Y_{1,organelle} + \frac{p_{unbound,ionised\ 2}}{p_{unbound,unionised}} \frac{N_2}{e^{N_2-1}} Y_{2,organelle}} \quad (16)$$

Where P is the permeability coefficient of either the ionised or unionised moiety. Using a ratio of the permeability coefficient between the unionised and ionised species we can describe the differential permeability of the two molecular forms. It has previously been assumed that the permeability coefficient of the ionised species is 3-4 log units lower than that of the neutral form (Trapp et al., 2008). Here we assume ionised, unionised permeability coefficient ratios of 2.3 and 3.3 log units for the monoprotic cation and monoprotic anion, respectively; for both diprotic cations and anions we assume a ratio of 10 log units; for ampholytes we assume a ratio of 2 log units.

Table 2. Nernst-Planck calculations

Compound type	N
Neutral	$N_1 = 0, N_2 = 0$
Monoprotic acid	$N_1 = -\frac{\phi F}{RT}, N_2 = 0$
Monoprotic base	$N_1 = \frac{\phi F}{RT}, N_2 = 0$
Diprotic acid	$N_1 = -\frac{\phi F}{RT}, N_2 = -\frac{\phi F}{RT}$
Diprotic base	$N_1 = \frac{\phi F}{RT}, N_2 = \frac{\phi F}{RT}$
Ampholyte	$N_{1acid} = -\frac{\phi F}{RT}, N_{2base} = \frac{\phi F}{RT}, N_2 = \frac{z\phi F}{RT} (z = 0)$

$$K_{cell,u} = \frac{C_{cell}}{C_{media_{unbound}}} = \left((1 - f_{lyso} + f_{mito}) \left(\frac{f_{iw}X_{iw} + P_{nl}f_{nl} + P_{np}f_{np}}{+K_{aAP}[AP^-](Y_{1,iw} + Y_{2,iw})} \right) + f_{lyso} \left(\frac{f_{iw}X_{lyso} + P_{nl}f_{nl} + P_{np}f_{np}}{K_{aAP}[AP^-](Y_{1,lyso} + Y_{2,lyso})} \right) K_{iw\ lyso}^{uu,uu} + f_{mito} \left(\frac{f_{iw}X_{mito} + P_{nl}f_{nl} + P_{np}f_{np}}{K_{aAP}[AP^-](Y_{1,mito} + Y_{2,mito})} \right) K_{iw\ mito}^{uu,uu} \right) \frac{1}{X_{ew}} K_{cell\ uu,uu} \quad (17)$$

Eq. 17 is adapted from the original Rodgers and Rowland (Rodgers et al., 2005; Rodgers and Rowland, 2007) approach incorporating $K_{cell\ uu,uu}$ where f_{iw} , f_{lyso} , f_{mito} , f_{nl} , f_{np} denote the fractional cellular volumes of the intracellular water, lysosomes, mitochondria, neutral lipids, and neutral phospholipids, respectively. P_{nl} and P_{np} describe the partitioning of the compound between intracellular water and neutral lipids and neutral phospholipids, respectively. Where the olive oil to water ($P_{vo:w}$) and octanol to water ($P_{o:w}$) partition coefficients are used as surrogates, respectively. K_{aAP} (μM^{-1}) is the acidic phospholipid association constant and $[AP^-]$ is the intracellular acidic phospholipid concentration. This may be determined experimentally for individual compounds or estimated from the blood to plasma ($B:P$) ratio using Eq. 18 (per Rodgers and Rowland (Rodgers et al., 2005; Rodgers and Rowland, 2007)).

$$K_{aAP} = \frac{\left[\frac{\frac{haem-1+B:P}{haem}}{fu} \cdot \frac{X_{plasma}}{K_{RBC\ uu,uu}} - (f_{iwRBC} \cdot X_{IW,RBC} + P_{nl} \cdot f_{nl,RBC} + P_{np} \cdot f_{np,RBC}) \right]}{\frac{1}{[AP^-]_{RBC} \cdot (Y_{1,RBC} + Y_{2,RBC})}} \quad (18)$$

As with Henry's law constant, experimentally determined $P_{o:w}$ values, and predicted values, are reported at a reference temperature of 25°C and these too must be corrected to correspond to the relevant system temperature before being used in as model inputs, Eq. 19; where ΔU_{ow} ($J\ mol^{-1}$) is the internal energy of phase change for octanol to water partitioning.

$$\log P_{ow,Tsys} = \log P_{ow} - \frac{\Delta U_{ow}}{\ln 10 \cdot R} \cdot \left(\frac{1}{T_{sys}} - \frac{1}{T_{ref}} \right) \quad (19)$$

As well as the incorporation of $K_{cell\ uu,uu}$, Eq. 17 has two further modifications beyond the original Rodgers and Rowland approach. Firstly, there is no longer a need to incorporate the fraction of extracellular water in tissues or binding to extracellular proteins present in the extracellular water since this fraction will be absent in 2D cell culture systems. Secondly, the equation is adapted so that two subcellular organelle compartments (lysosome and mitochondria) are considered.

While Eq. 17 allows the calculation of total intracellular concentrations, it is necessary to determine separate cellular compartment to medium ratios to be able to calculate the concentration within respective cellular compartment; for details see the *supplementary material*.

2.7. Apparent In Vitro Volume of Distribution at Steady-State

Using the partitioning coefficients, serum-binding and system parameters described above, we can predict the apparent *in vitro* volume of distribution at steady-state, Eq. 20, thus generating a prediction of the unbound, freely dissolved media concentration.

$$C_{medium,dissolved,u} = \frac{C_{nominal} \cdot f_{uFBS,dilu} \cdot V_{medium}}{V_{medium} + K_{air} f_{ui} V_{air} + K_{cell} V_{totalcell} + K_{plastic} S A_{media} \cdot 10^3} \quad (20)$$

Subsequently, the steady-state concentration in each compartment of the model can be predicted based on the respective partition coefficient and the unbound, freely dissolved media concentration (Eq. 21, Eq. 22, Eq. 23).

$$C_{air} = K_{air,u} \cdot f_{ui} \cdot C_{medium,dissolved,u} \quad (21)$$

$$\text{if } V_{air} = 0, C_{air} = 0$$

$$C_{plastic} = K_{plastic,u} \cdot C_{medium,dissolved,u} \cdot 10^3 \quad (22)$$

$$C_{cell} = K_{cell,u} \cdot C_{medium,dissolved,u} \quad (23)$$

2.8. Quasi Steady-State Dynamic modelling

The partition model presented defines the system at steady-state. To test the potential impact of the steady-state assumption, we formulated a dynamic variant of the model incorporating the instantaneous rate of metabolism or loss from the system by the following differential equation, Eq. 24.

$$\frac{dQ_{met}(t)}{dt} = n_{cell} \cdot CL_{int,met} \cdot C_{cell}(t) \quad (24)$$

Where $Q_{met}(t)$ is the quantity metabolised in moles up to time t , $CL_{int,met}$ (L min⁻¹) is the intrinsic metabolic clearance per cell, and n_{cell} is the number of cells in the cultured monolayer. The total quantity of parent chemical remaining in the system at time t , $Q_{total}(t)$, can then be calculated, Eq. 25.

$$Q_{total}(t) = C_{nominal} \cdot V_{media} - Q_{met}(t) \quad (25)$$

In Eq. 24, $C_{cell}(t)$ is computed using Eq. 20 and Eq. 23, replacing $C_{nominal}$ by $\frac{Q_{total}(t)}{V_{media}}$ at each integration time step; Eq. 21 to Eq. 23 apply without change. This quasi-steady-state differential formulation is flexible and could be easily modified to describe nonlinear metabolism or other *in vitro* elimination processes if needed. In its current linear form, its

analytical solution is known and could be used. However, for the simulation results presented in the following we numerically integrated the differential equation above.

2.9. Code and simulations

The static model equations have been coded in R (R Core Team, 2018), and GNU MCSim (Bois, 2009) for the dynamic model.

3. Results

3.1. Steady-state model

We performed representative simulations of the steady-state distribution of 1194 compounds with diverse physicochemical properties in an adherent monolayer of primary human hepatocytes cultured in 10% FBS assuming a 1nM nominal treatment concentration similar to Armitage and colleagues (Armitage et al., 2014). Firstly, steady-state distribution was simulated assuming no significant ionisation of test compounds; under this assumption the VIVD model shows largely comparable predictions of total intracellular to that previously published (Armitage et al., 2014) (Figure 1).

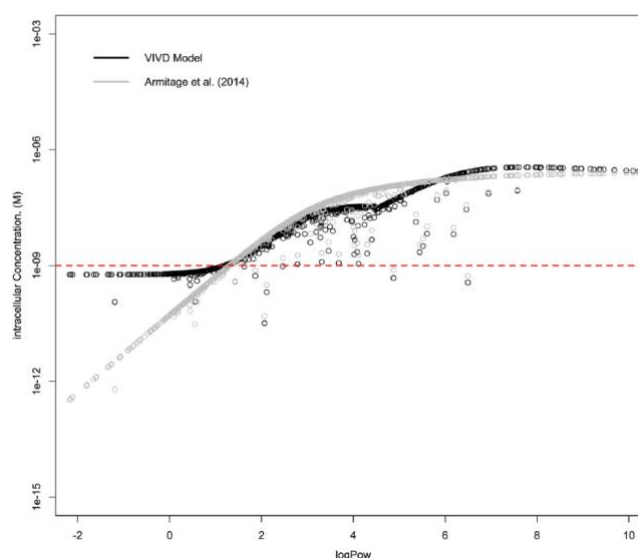


Figure 1. Comparative simulations of steady-state distribution *in vitro* with the previously published (9) (grey) and VIVD model (black) assuming neutrality; simulations represent primary human hepatocytes cultured in 10% FBS containing 24 and 1.9 gL⁻¹ of albumin and TAG, respectively treated with 1nM (red dotted line) of each test compound individually.

This is to be expected given the underlying assumptions of the respective models; here we describe the distribution into cells based on the partitioning into the intracellular water as

well as the binding to cellular neutral lipids and phospholipids. Although our model does account for binding to cellular acidic phospholipids, as we assume that this is only relevant for significantly ionised basic compounds and so does not contribute to the intracellular binding of neutral compounds. The Armitage model (Armitage et al., 2014) defines partitioning into cells based on the fraction of the cell comprised of lipids, f_{lipid} , and the octanol to water partition coefficient (Eq. 26).

$$K_{cell} = f_{lipid}P_{ow} \quad (26)$$

The total fraction of the cell comprised of neutral lipids and phospholipids in our model is very similar to that in simulations with the Armitage model (0.06 versus 0.05, respectively) and so for neutral compounds we would expect comparable predictions. However, like the model recently published by Fischer *et al.* (Fischer et al., 2017), our VIVD model also incorporates the partitioning of test compound between the medium water and intracellular water and so predicts higher intracellular concentrations at $\log P_{ow} < 1$.

Simulating the same cell culture system, but incorporating simulated, randomly assigned, differential ionisation of test compounds, rather than the assumption of neutrality, we see significant differences in the predicted total intracellular concentrations (Figure 2). Moderate to strong bases ($pK_a \geq 7$) show significantly higher partitioning into cells than would be predicted assuming neutrality. In part this is attributable to the differential ionisation of basic compounds between the culture medium ($pH=7.4$) and the intracellular water ($pH=7.0$). The fraction ionised in the intracellular water exceeds that in the culture medium, with this fraction having reduced membrane permeability, relative to the neutral fraction, resulting in an increased proportion of the free dose entering and remaining within the cells. Additionally, while the ionised fraction of compound within the culture media has reduced membrane permeability, the negative membrane across the hepatocyte cell membrane ($\phi=-70mV$) increases the apparent permeability of ionised (cationic) bases. Conversely, significantly ionised acids show significantly reduced predicted partitioning into cells when ion permeability and the impact of membrane potential are considered within the model, due to ionisation in the culture media and the limiting effect of the negative membrane potential on the ionised (anionic) fraction. The total intracellular concentrations predicted under the two different model assumptions can differ by many orders of magnitude and so could significantly impact the *in vitro* to *in vivo* extrapolation of toxicity endpoints quantified *in vitro*.

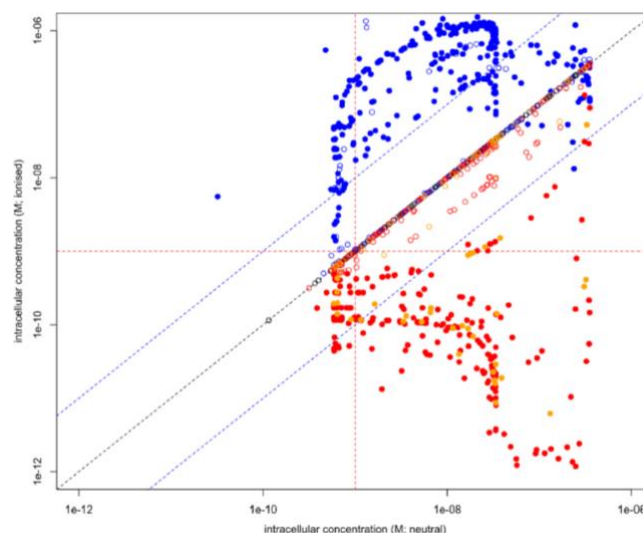


Figure 2. Comparative simulations of steady-state distribution in vitro with the VIVD model assuming neutrality (x-axis) and accounting for ionisation (y-axis); simulations represent primary human hepatocytes cultured in 10% FBS containing 24 and 1.9 gL^{-1} of albumin and TAG, respectively treated with 1nM (red dotted line) of each test compound individually; acids (red), bases (blue), ampholytes (orange). The black dotted line represents unity, blue dotted lines represent 10-fold, compounds with fraction unionised <0.1 (pH=7) plotted as solid data points.

The prediction of binding to components within serum is predominantly driven by $\log P_{o:w}$ even with accounting for ionisation within the model framework. Binding to serum albumin is the most significant contributor to $f_{u_{FBS}}$ and, since the model assumes that ionisation does not impact a compound's ability to bind to albumin, the assumption that only the unionised fraction of test compound binds to serum triglycerides has a minimal effect on overall serum binding in our simulations (Figure 3). Although the prediction of $f_{u_{FBS}}$ accounts for ionisation, it does not differentiate between ionised species (i.e. acids vs bases).

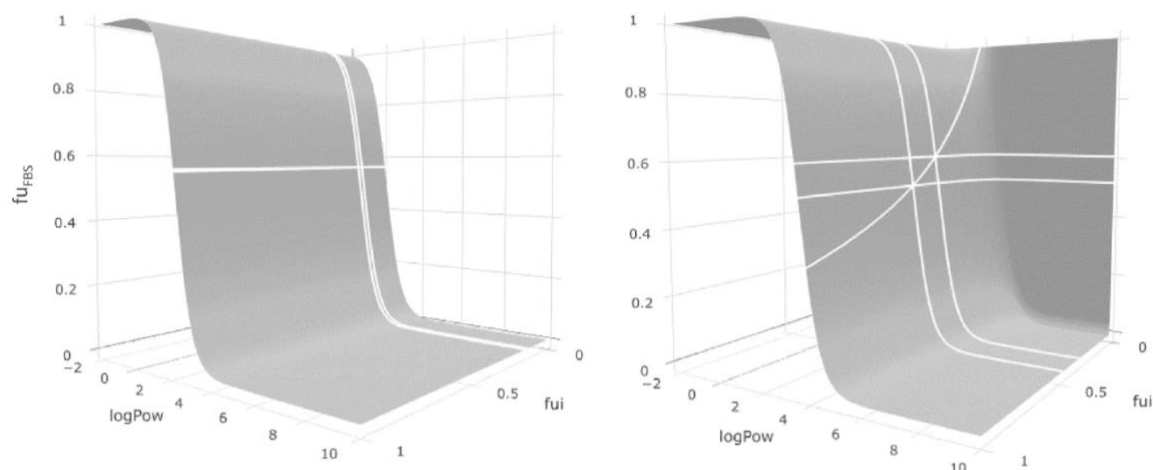


Figure 3. Predicted fraction unbound in fetal bovine; serum containing 24 and 1.9 gL⁻¹ of albumin and TAG, respectively (left), serum containing 0 and 1.9 gL⁻¹ of albumin and TAG, respectively (right).

3.2. Quasi steady-state dynamic model

For 612 of the 1194 chemicals considered (Armitage et al., 2014), we have used hepatic metabolic clearance estimates reported in the *R* package *httk* (Wambaugh et al., 2018) (v1.7). Those estimates were corrected for the fraction unbound in medium by Wambaugh (Wambaugh et al., 2018) and colleagues, but not for the partitioning into cells (i.e. for the partition coefficient $K_{cell,u}$) so we divided the values given in *httk* by that parameter, correcting units as required. For those 612 chemicals, using our dynamic model, we performed simulations of the time courses over 48h for the quantity of parent chemical remaining in the system and for the intracellular cellular concentration, following a single treatment at a nominal concentration of 1 nM, Figure 4. For the majority of the 612 chemicals, metabolism has an impact; 283 have been more than 90% metabolized and 165 of them have been between 10 and 90% metabolized. Under the quasi-steady-state assumptions, the concentrations in cells, and other *in vitro* compartments are affected to the same proportion. This demonstrates that the steady-state model would over-predict intracellular exposure in metabolically competent *in vitro* cell systems for the majority of the 612 subsets of compounds.

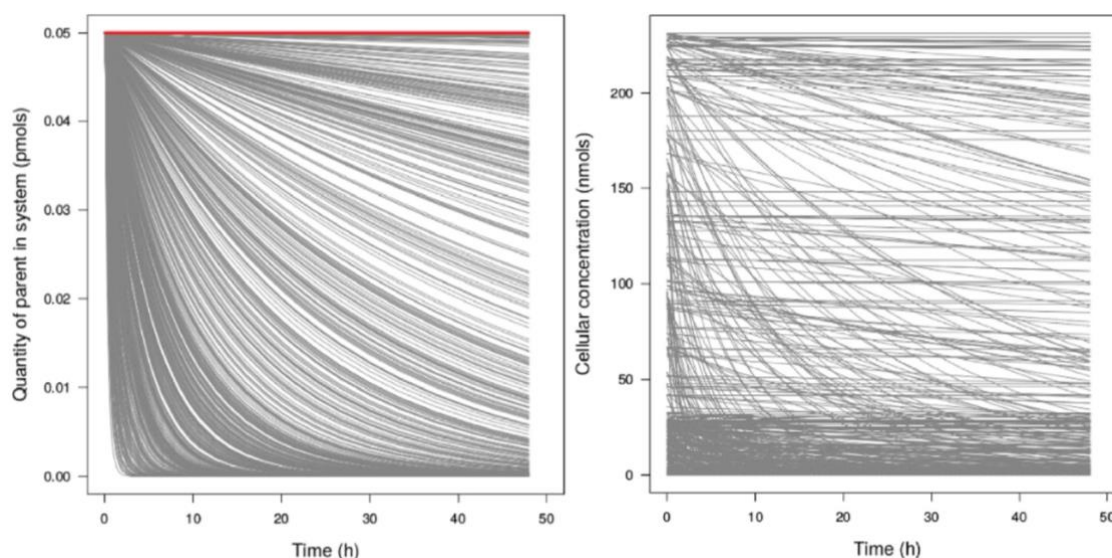


Figure 4. Predictions of the quantity-time profile of parent chemical in the assay system (left) and of its intracellular concentration-time profile (right); predictions assuming no metabolism (red line) and incorporating metabolism (grey lines) for 612 chemicals.

4. Discussion

Here we present a pragmatic modelling approach for predicting the distribution of test compounds in *in vitro* cell culture systems. The aim is to predict the intracellular concentration accounting for binding to serum components, culture plastics, and accounting for distribution into the air in the headspace above the culture media, and binding to cellular lipids. The approach detailed above expands on previously published approaches (Armitage et al., 2014; Fischer et al., 2017; Zaldivar Comenges et al., 2017) by incorporating differential ionisation of test compounds between the intracellular water and the aqueous culture media, ion-permeability across cell membranes, and the effect of membrane potential. Our results show that predicted intracellular concentrations differ considerably dependent on whether the model used disregards the impact of compound ionisation (i.e. all compounds are neutral) or accounts for compound ionisation. While the partitioning of compound between culture medium and cells in our model is based upon the previously published work of Rodgers and Rowland (Rodgers et al., 2005; Rodgers and Rowland, 2007), we revise a fundamental assumption incorporating the permeability of the ionised species, rather than assuming that only the unionised form will partition across the membrane.

With the exception of toxicities mediated through targets presented at the cell surface, the unbound intracellular concentration is the most directly relevant operating concentration

driving toxicological endpoints observed *in vitro*, and therefore the most translatable to *in vivo* scenarios. Previous publications have sought to improve on the IVIVE of concentration-dependent toxicity by correcting for binding to serum-supplemented culture media, so translating the unbound effect concentration *in vitro* to the unbound plasma concentration *in vivo* (Gulden and Seibert, 2003). External *in vivo* doses required to achieve the unbound effect-concentration, determined *in vitro* can then be predicted based upon reverse-dosimetry using physiologically-based pharmacokinetic (PBPK) models. However, while representative at many levels of their tissue of origin, isolated primary cells or immortalised/transformed cell lines are very different to *in vivo* tissues lacking key structural elements (e.g. interstitial fluid). As such, the most translatable operating concentration from *in vitro* is the intracellular concentration, which can inform a reverse-dosimetry approach using PBPK models describing the distribution of compounds into tissues, incorporating extracellular water and tissue blood flow.

A balance must be struck between the pragmatic approach presented here, and more complex models (Zaldivar Comenges et al., 2017). The model presented here predicts *in vitro* distribution at steady-state; while this is an acceptable approximation for predictions in non-proliferating cell systems with low metabolic capacity or stable compounds, it could be misleading if these model assumptions are in conflict with the reality of the cell culture system being used. This would be important for example if trying to predict intracellular concentrations in 3D culture systems, hanging droplet or spheroid systems, or organ on a chip/micro-physiological system. We demonstrate the impact of metabolic clearance on predictions relative to the steady-state model with a quasi-steady-state dynamic version of the model that is applicable in metabolically competent cellular systems, such as primary human hepatocytes. The subset of 612 chemicals used for the dynamic system simulations cannot be taken as a representative sample of all chemicals. However, it demonstrates that in a metabolically competent assay system, metabolism can have a large impact on intracellular exposure. In that case, the dynamic version of the model should be used to obtain more accurate estimates of effective dose. As noted above, the quasi-steady-state differential formulation is flexible and could describe other *in vitro* elimination processes if needed. Such an adaption should be considered when simulating the distribution of highly volatile compounds in *in vitro* assays. In a non-hermetically sealed system, the loss of compound to the air represents a clearance mechanism that would lead to over-prediction of intracellular concentrations by the model under steady-state assumptions. Outputs from the steady-state model showing significant

distribution of test compound into the headspace, will identify the need for users to consider whether the model assumptions are fit for purpose or necessitate revision. It may also identify the need for further considerations in laboratory safety when performing assays, and so may inform protocol design.

5. Conclusion

In vitro toxicity assays are increasingly being used to identify chemical hazards to human and animal health and inform risk assessment. Consequently, these systems must be well characterised, and the kinetics understood in order to extract the correct parameters and better interpret the toxicity markers monitored *in vitro*. Ultimately, such an approach will better inform quantitative *in vitro* to *in vivo* extrapolation and elucidate mechanistic insight as part of a systems toxicology approach (Hartung et al., 2017). While only limited verification of such models has been published to date, and further experimental work are warranted (Bellwon et al., 2015; Kramer et al., 2012), model development and verification will need to be performed in tandem to the development of increasingly sophisticated *in vitro* systems in order to assess the impact of *in vitro* distribution and facilitate the robust IVIVE of chemical hazard (Pamies and Hartung, 2017).

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 681002.

References

- Armitage, J.M., Wania, F., Arnot, J.A., 2014. Application of mass balance models and the chemical activity concept to facilitate the use of *in vitro* toxicity data for risk assessment. *Environ Sci Technol* 48, 9770-9779.
- Bell, S.M., Chang, X., Wambaugh, J.F., Allen, D.G., Bartels, M., Brouwer, K.L.R., Casey, W.M., Choksi, N., Ferguson, S.S., Fraczekiewicz, G., Jarabek, A.M., Ke, A., Lumen, A., Lynn, S.G., Paini, A., Price, P.S., Ring, C., Simon, T.W., Sipes, N.S., Sprankle, C.S., Strickland, J., Troutman, J., Wetmore, B.A., Kleinstreuer, N.C., 2018. *In vitro* to *in vivo* extrapolation for high throughput prioritization and decision making. *Toxicol In Vitro* 47, 213-227.
- Bellwon, P., Truissi, G.L., Bois, F.Y., Wilmes, A., Schmidt, T., Savary, C.C., Parmentier, C., Hewitt, P.G., Schmal, O., Josse, R., Richert, L., Guillouzo, A., Mueller, S.O., Jennings, P., Testai, E., Dekant, W., 2015. Kinetics and dynamics of cyclosporine A in three hepatic cell culture systems. *Toxicol In Vitro* 30, 62-78.
- Berezhkovskiy, L.M., 2011. The corrected traditional equations for calculation of hepatic clearance that account for the difference in drug ionization in extracellular and intracellular tissue water and the corresponding corrected PBPK equation. *J Pharm Sci* 100, 1167-1183.

- Blaauboer, B.J., 2010. Biokinetic modeling and in vitro-in vivo extrapolations. *J Toxicol Environ Health B Crit Rev* 13, 242-252.
- Bois, F.Y., 2009. GNU MCSim: Bayesian statistical inference for SBML-coded systems biology models. *Bioinformatics* 25, 1453-1454.
- Dearden, J.C., Schuurmann, G., 2003. Quantitative structure-property relationships for predicting Henry's law constant from molecular structure. *Environ Toxicol Chem* 22, 1755-1770.
- Deckelbaum, R.J., Granot, E., Oschry, Y., Rose, L., Eisenberg, S., 1984. Plasma triglyceride determines structure-composition in low and high density lipoproteins. *Arteriosclerosis* 4, 225-231.
- Endo, S., Goss, K.U., 2011. Serum albumin binding of structurally diverse neutral organic compounds: data and models. *Chem Res Toxicol* 24, 2293-2301.
- Fischer, F.C., Henneberger, L., Konig, M., Bittermann, K., Linden, L., Goss, K.U., Escher, B.I., 2017. Modeling Exposure in the Tox21 in Vitro Bioassays. *Chem Res Toxicol* 30, 1197-1208.
- Ghosh, A., Maurer, T.S., Litchfield, J.E., Varma, M.V., Rotter, C.J., Scialis, R., Feng, B., Tu, M., Guimaraes, C.R.W., Scott, D.O., 2014a. Towards a Unified Model of Passive Drug Permeation II: The Physiochemical Determinants of Unbound Tissue Distribution with Applications to the Design of Hepatoselective Glucokinase Activators. *Drug Metab Dispos* 42, 1599-1610.
- Ghosh, A., Scott, D.O., Maurer, T.S., 2014b. Towards a unified model of passive drug permeation I: origins of the unstirred water layer with applications to ionic permeation. *Eur J Pharm Sci* 52, 109-124.
- Gibb, S., 2008. Toxicity testing in the 21st century: a vision and a strategy. *Reprod Toxicol* 25, 136-138.
- Godoy, P., Hewitt, N.J., Albrecht, U., Andersen, M.E., Ansari, N., Bhattacharya, S., Bode, J.G., Bolleyn, J., Borner, C., Bottger, J., Braeuning, A., Budinsky, R.A., Burkhardt, B., Cameron, N.R., Camussi, G., Cho, C.S., Choi, Y.J., Craig Rowlands, J., Dahmen, U., Damm, G., Dirsch, O., Donato, M.T., Dong, J., Dooley, S., Drasdo, D., Eakins, R., Ferreira, K.S., Fonsato, V., Fraczek, J., Gebhardt, R., Gibson, A., Glanemann, M., Goldring, C.E., Gomez-Lechon, M.J., Groothuis, G.M., Gustavsson, L., Guyot, C., Hallifax, D., Hammad, S., Hayward, A., Haussinger, D., Hellerbrand, C., Hewitt, P., Hoehme, S., Holzhutter, H.G., Houston, J.B., Hrach, J., Ito, K., Jaeschke, H., Keitel, V., Kelm, J.M., Kevin Park, B., Kordes, C., Kullak-Ublick, G.A., LeCluyse, E.L., Lu, P., Luebke-Wheeler, J., Lutz, A., Maltman, D.J., Matz-Soja, M., McMullen, P., Merfort, I., Messner, S., Meyer, C., Mwinyi, J., Naisbitt, D.J., Nussler, A.K., Olinga, P., Pampaloni, F., Pi, J., Pluta, L., Przyborski, S.A., Ramachandran, A., Rogiers, V., Rowe, C., Schelcher, C., Schmich, K., Schwarz, M., Singh, B., Stelzer, E.H., Stieger, B., Stober, R., Sugiyama, Y., Tetta, C., Thasler, W.E., Vanhaecke, T., Vinken, M., Weiss, T.S., Widera, A., Woods, C.G., Xu, J.J., Yarborough, K.M., Hengstler, J.G., 2013. Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. *Arch Toxicol* 87, 1315-1530.
- Gstraunthaler, G., 2003. Alternatives to the use of fetal bovine serum: serum-free cell culture. *ALTEX* 20, 275-281.

- Gulden, M., Seibert, H., 2003. In vitro-in vivo extrapolation: estimation of human serum concentrations of chemicals equivalent to cytotoxic concentrations in vitro. *Toxicology* 189, 211-222.
- Hartung, T., FitzGerald, R.E., Jennings, P., Mirams, G.R., Peitsch, M.C., Rostami-Hodjegan, A., Shah, I., Wilks, M.F., Sturla, S.J., 2017. Systems Toxicology: Real World Applications and Opportunities. *Chem Res Toxicol* 30, 870-882.
- Hoehme, S., Brulport, M., Bauer, A., Bedawy, E., Schormann, W., Hermes, M., Puppe, V., Gebhardt, R., Zellmer, S., Schwarz, M., Bockamp, E., Timmel, T., Hengstler, J.G., Drasdo, D., 2010. Prediction and validation of cell alignment along microvessels as order principle to restore tissue architecture in liver regeneration. *Proc Natl Acad Sci USA* 107, 10371-10376.
- Holmes, A.M., Creton, S., Chapman, K., 2010. Working in partnership to advance the 3Rs in toxicity testing. *Toxicology* 267, 14-19.
- Katritzky, A.R., Wang, Y., Sild, S., Tamm, T., 1998. QSPR Studies on Vapor Pressure, Aqueous Solubility, and the Prediction of Water-Air Partition Coefficients. *J Chem Inf Comput Sci* 38, 720-725.
- Kazmi, F., Hensley, T., Pope, C., Funk, R.S., Loewen, G.J., Buckley, D.B., Parkinson, A., 2013. Lysosomal sequestration (trapping) of lipophilic amine (cationic amphiphilic) drugs in immortalized human hepatocytes (Fa2N-4 cells). *Drug Metab Dispos* 41, 897-905.
- Kramer, N.I., 2010. Measuring, modelling and increasing the free concentration of test chemicals in cell assays. PhD Thesis, Utrecht University, Utrecht, Netherlands.
- Kramer, N.I., Krismartina, M., Rico-Rico, A., Blaauboer, B.J., Hermens, J.L., 2012. Quantifying processes determining the free concentration of phenanthrene in Basal cytotoxicity assays. *Chem Res Toxicol* 25, 436-445.
- Kupke, D.W., Hodgins, M.G., Beams, J.W., 1972. Simultaneous determination of viscosity and density of protein solutions by magnetic suspension. *Proc Natl Acad Sci USA* 69, 2258-2262.
- Leist, M., Ghallab, A., Graepel, R., Marchan, R., Hassan, R., Bennekou, S.H., Limonciel, A., Vinken, M., Schildknecht, S., Waldmann, T., Danen, E., van Ravenzwaay, B., Kamp, H., Gardner, I., Godoy, P., Bois, F.Y., Braeuning, A., Reif, R., Oesch, F., Drasdo, D., Hohme, S., Schwarz, M., Hartung, T., Braunbeck, T., Beltman, J., Vrieling, H., Sanz, F., Forsby, A., Gadaleta, D., Fisher, C., Kelm, J., Fluri, D., Ecker, G., Zdrazil, B., Terron, A., Jennings, P., van der Burg, B., Dooley, S., Meijer, A.H., Willighagen, E., Martens, M., Evelo, C., Mombelli, E., Taboureaux, O., Mantovani, A., Hardy, B., Koch, B., Escher, S., van Thriel, C., Cadenas, C., Kroese, D., van de Water, B., Hengstler, J.G., 2017. Adverse outcome pathways: opportunities, limitations and open questions. *Arch Toxicol* 91, 3477-3505.
- Meek, M.E., Lipscomb, J.C., 2015. Gaining acceptance for the use of in vitro toxicity assays and QIVIVE in regulatory risk assessment. *Toxicology* 332, 112-123.
- Pamies, D., Hartung, T., 2017. 21st Century Cell Culture for 21st Century Toxicology. *Chem Res Toxicol* 30, 43-52.
- R Core Team, 2018. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.
- Rodgers, T., Leahy, D., Rowland, M., 2005. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. *J Pharm Sci* 94, 1259-1276.

- Rodgers, T., Rowland, M., 2007. Physiologically-based Pharmacokinetic Modeling 2: Predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. *J Pharm Sci* 95, 1238-1257.
- Schwarzenbach, R.P., Gschwen, P.M., Imoden, D.M., 2004. *Environmental Organic Chemistry*. J. Wiley and Sons.
- Schwen, L.O., Schenk, A., Kreutz, C., Timmer, J., Bartolome Rodriguez, M.M., Kuepfer, L., Preusser, T., 2015. Representative Sinusoids for Hepatic Four-Scale Pharmacokinetics Simulations. *PLoS One* 10, e0133653.
- Tice, R.R., Austin, C.P., Kavlock, R.J., Bucher, J.R., 2013. Improving the human hazard characterization of chemicals: a Tox21 update. *Environ Health Perspect* 121, 756-765.
- Trapp, S., Rosania, G.R., Horobin, R.W., Kornhuber, J., 2008. Quantitative modeling of selective lysosomal targeting for drug design. *Eur Biophys J* 37, 1317-1328.
- Wambaugh, J.F., Hughes, M.F., Ring, C.L., MacMillan, D.K., Ford, J., Fennell, T.R., Black, S.R., Snyder, R.W., Sipes, N.S., Wetmore, B.A., Westerhout, J., Setzer, R.W., Pearce, R.G., Simmons, J.E., Thomas, R.S., 2018. Evaluating In Vitro-In Vivo Extrapolation of Toxicokinetics. *Toxicol Sci* 163, 152-169.
- Worth, A.P., Louisse, J., Macko, P., Sala Benito, J.V., Paini, A., 2017. Virtual Cell Based Assay simulations of intra-mitochondrial concentrations in hepatocytes and cardiomyocytes. *Toxicol In Vitro* 45, 222-232.
- Zaldivar Comenges, J.M., Joossens, E., Benito, J.V.S., Worth, A., Paini, A., 2017. Theoretical and mathematical foundation of the Virtual Cell Based Assay - A review. *Toxicol In Vitro* 45, 209-221.
- Zheng, X., Baker, H., Hancock, W.S., Fawaz, F., McCaman, M., Pungor, E., Jr., 2006. Proteomic analysis for the assessment of different lots of fetal bovine serum as a raw material for cell culture. Part IV. Application of proteomics to the manufacture of biological drugs. *Biotechnol Prog* 22, 1294-1300.

CHAPITRE III. Modèle PBPK pour évaluer de la toxicité développementale chez l'embryon de poisson zèbre

1. Un nouveau modèle PBPK générique pour l'embryon de poisson zèbre

1.1. Objectif des travaux et contribution

Pour diverses raisons exposées dans le Chapitre 1, l'embryon de poisson zèbre est particulièrement attractif en toxicologie et pharmacologie. L'extrapolation de la toxicité du poisson zèbre à l'homme nécessite, cependant de tenir compte des différences pharmacocinétiques entre les deux espèces. De plus, la connaissance des concentrations chimiques dans les organes est fondamentale pour comprendre les relations dose-réponse. Les concentrations internes sont souvent difficiles à mesurer, notamment chez des espèces aussi petites que l'embryon de poisson zèbre. Cependant, en reliant l'anatomie, la physiologie et les processus biochimiques pour comprendre et calculer le comportement des composés au sein d'un organisme, les modèles PBPK permettent de prédire approximativement des profils cinétiques de concentration en fonction du temps dans des organes expérimentalement inaccessibles, à partir de données expérimentales. Dans cette optique, nous avons développé un modèle PBPK générique sous forme d'équations différentielles et par approche bayésienne, intégrant la croissance des organes et le métabolisme hépatique. Le modèle suppose une distribution à l'état de quasi équilibre entre les cellules, les lysosomes et les mitochondries dans différents tissus de l'embryon (sac vitellin, foie, intestin, muscle, squelette, œil, cerveau, cœur, peau et autres tissus). Le modèle est générique dans la mesure où il peut simuler la distribution de nombreux produits chimiques sur la base de leurs propriétés physico-chimiques : les coefficients de partage des produits chimiques entre les cellules ou organelles et le milieu de culture sont prédits via le modèle présenté précédemment dans le Chapitre 2 de la distribution intracellulaire *in vitro* virtuelle. Pour atténuer l'incertitude due à leur prédiction, nous leur avons attribué un facteur de correction.

Nous avons appliqué notre modèle à l'analyse des données de toxicité développementale de composés cibles du projet EU-ToxRisk, à savoir, l'acide valproïque et ses analogues : acide 2,2-diméthylvalérique, acide 2-éthylbutyrique, acide 2-éthylhexanoïque, acide 2-méthylhexanoïque, acide 2-méthylpentanoïque, Acide 2-propylheptanoïque, 4-èneVPA, acide

4-penténoïque et acide hexanoïque. Les paramètres estimés pour améliorer la prédictivité de notre modèle sont la clairance métabolique et le facteur de correction des coefficients de partage. Ils ont été estimés simultanément par ajustement bayésien aux données de concentrations de VPA et de ses analogues dans l'embryon et le milieu de culture au cours de cinq jours de dosages répétés. Enfin, pour prétendre à des extrapolations inter-espèces pertinentes et pour améliorer la compréhension de la relation dose-réponse, nous avons déterminé les concentrations effectives de 10% de réponse (EC₁₀) sur la base des concentrations totales embryonnaires prédites plutôt que des concentrations nominales. Par nos travaux, nous démontrons l'utilisation du modèle pour classer les composés selon leur potentiel toxique à partir d'estimation de concentrations internes plutôt que de concentrations nominales, pour ainsi prétendre à l'extrapolation des résultats des tests d'embryons de poisson zèbre à l'évaluation des risques humains.

Ces travaux de recherche ont été valorisés dans un article « Development of a generic zebrafish embryo PBPK model and application to the developmental toxicity assessment of valproic acid analogs », présenté ci-après et publié dans la revue Reproductive Toxicology en 2020 (DOI : 10.1016/j.reprotox.2020.02.010). Les informations supplémentaires sont disponibles en Annexe 2. De plus, le cas d'étude du projet européen EU-ToxRisk intégrant nos travaux fera prochainement l'objet d'une publication dans le journal Alternatives to animal experimentation.

1.2. Articulation avec le modèle VIVD précédent

Les tests sur embryons de poisson zèbre sont assimilables en partie à un système cellulaire *in vitro* car le protocole de l'OCDE présenté dans le chapitre 1 préconise de placer les embryons, avec de l'eau, dans les puits d'une plaque de culture cellulaire multi-puits. Comme dans tout système *in vitro*, les molécules présentes dans le milieu de culture se distribuent dans l'embryon, ses divers organes, l'air, le plastique et autres composants potentiellement présents, en fonction de leurs propriétés physico-chimiques. On peut donc utiliser le modèle de distribution virtuelle *in vitro* pour déterminer les coefficients de partage entre le milieu de culture et les différents types cellulaires de l'embryon (en fonction de la composition spécifique des cellules), mais également la liaison au plastique et la capacité d'évaporation, pour les composés analysés, à savoir l'acide valproïque et neuf de ses analogues. Ces coefficients de partage font partie des paramètres d'entrée du modèle PBPK.

1.3. Nécessité d'un corrigé

L'article est ici suivi d'un corrigé (sous presse dans la revue Reproductive Toxicology). Il s'est avéré, lors de la revue de l'ensemble des résultats de la deuxième étude de cas de EU-ToxRisk, que l'eau dans laquelle baignaient les embryons n'était pas suffisamment tamponnée pour maintenir un pH constant après l'ajout des différentes concentrations d'acides organiques (VPA et analogues) testés. Nous avons fait l'hypothèse que le pH était maintenu à 7.4, mais une variation selon la concentration a été observée. Le problème est que le pH du milieu de culture est un paramètre essentiel du modèle VIVD car il conditionne le degré d'ionisation des molécules testées et donc leur capacité à s'accumuler dans l'embryon. Les pH actuellement vus par les embryons au cours des expérimentations ont été mesurés *a posteriori* pour chaque concentration testée et nous avons donc refait tous les calculs publiés en tenant compte de la valeur correcte de pH.

1.4. Perspectives

Le modèle PBPK embryon de poisson zèbre peut être appliqué pour répondre à différentes problématiques (éco)-toxicologiques. Dans un premier temps, par exploitation des données physico-chimiques et des prédictions QSAR du métabolisme, il peut être utilisé comme outil *ab initio* à haut débit pour prédire approximativement la distribution des composés dans différents tissus embryonnaires, en fonction du temps et des concentrations d'exposition. Dans un second temps, si des données expérimentales de concentrations totales embryonnaires ou tissulaires sont disponibles, les paramètres du modèle peuvent être calibrés pour estimer le métabolisme *in vivo* et affiner les prédictions. Notre modèle peut également être utilisé pour orienter de nouvelles recherches et aider à concevoir de futures expériences. Par exemple, avec les données nécessaires, il pourrait être amélioré pour étudier l'impact du chorion ou des transporteurs sur la diffusion. De plus, les concentrations embryonnaires prédites permettent d'étudier plus justement la relation dose-réponse, si bien que nous avons développé le modèle multi-états markovien présenté dans le Chapitre 4 sur la base de ces concentrations internes plutôt que sur les concentrations nominales.

2. Résumé et article du modèle PBPK

Afin de mieux expliquer, prédire ou extrapoler à l'homme les effets de la toxicité développementale des produits chimiques sur les embryons de poisson zèbre (*Danio rerio*), nous avons développé un modèle de pharmacocinétique physiologique (PBPK) pour prédire les concentrations en composés neutres ou ionisables dans les différents organes, jusqu'à 120 heures après la fécondation. La distribution des composés est modélisée dans les cellules, les lysosomes et les mitochondries de dix organes de l'embryon. Les coefficients de partage du modèle sont calculés par l'intermédiaire de sous-modèles basés sur les propriétés physico-chimiques des produits chimiques d'intérêt. Le modèle considère la croissance des organes et les variations de la clairance métabolique au cours du temps. Nous avons comparé les prédictions du modèle *ab initio* aux données de concentrations du milieu de culture et embryonnaires d'acide valproïque (VPA) et de neuf analogues pendant l'administration continue, et selon la ligne directrice 236 de l'OCDE. Par calibration bayésienne, nous avons amélioré les prédictions en estimant la clairance métabolique et les coefficients de partage à partir des données. Nous avons également évalué la performance du modèle à reproduire les données publiées par (Brox et al., 2016b) du VPA et 16 autres produits chimiques. Nous avons finalement comparé les relations dose-réponse calculées pour la mortalité et les malformations sur la base des concentrations d'embryons entiers prédites par rapport à celles basées sur les concentrations nominales dans l'eau. L'utilisation de concentrations d'organes cibles a considérablement modifié l'ampleur des paramètres dose-réponse et le classement de la toxicité relative des produits chimiques étudiés.

Development of a generic zebrafish embryo PBPK model and application to the developmental toxicity assessment of valproic acid analogs[†]

Ségolène Siméon[†], Katharina Brotzmann[‡], Ciaran Fisher[§], Iain Gardner[§], Steve Silvester^{||}, Richard MacLennan^{||}, Paul Walker^{||}, Thomas Braunbeck[‡] and Frederic Y. Bois^{*,§}.

[†] INERIS, METO unit, Parc ALATA BP2, Verneuil en Halatte, France.

[‡] University of Heidelberg, Aquatic Ecology and Toxicology, Centre for Organismal Studies (COS), Im Neuenheimer Feld 504, D-69120 Heidelberg, Germany.

[§] CERTARA UK Limited, Simcyp Division, Level 2-Acero, 1 Concourse Way, Sheffield, S1 2BJ, United Kingdom; Email: *frederic.bois@certara.com

^{||} Cyprotex Discovery Ltd., No. 24 Mereside, Alderley Park, Macclesfield, Cheshire, SK10 4TG, United Kingdom.

Abstract

In order to better explain, predict, or extrapolate to humans the developmental toxicity effects of chemicals to zebrafish (*Danio rerio*) embryos, we developed a physiologically-based pharmacokinetic (PBPK) model designed to predict organ concentrations of neutral or ionizable chemicals, up to 120 hours post-fertilization. Chemicals' distribution is modeled in the cells, lysosomes, and mitochondria of ten organs of the embryo. The model's partition coefficients are calculated with sub-models using physicochemical properties of the chemicals of interest. The model accounts for organ growth and changes in metabolic clearance with time. We compared *ab initio* model predictions to data obtained on culture medium and embryo concentrations of valproic acid (VPA) and nine analogs during continuous dosing under the OECD test guideline 236. We further improved the predictions by estimating metabolic clearance and partition coefficients from the data by Bayesian calibration. We also assessed the performance of the model at reproducing data published by Brox et al. (2016) on VPA and 16 other chemicals. We finally compared dose-response relationships calculated for mortality and malformations on the basis of predicted whole embryo concentrations versus those based on nominal water concentrations. The use of target organ concentrations substantially shifted the magnitude of dose-response parameters and the relative toxicity ranking of chemicals studied.

Keywords

PBPK model, *Danio rerio*, Zebrafish embryo, Bayesian, Development, Effect concentration, Valproic acid, Internal concentration, Toxicity.

[†] Reproductive Toxicology 93, 219–229. <https://doi.org/10.1016/j.reprotox.2020.02.010>

1. Introduction

Prediction of chemicals' developmental and reproductive toxicity is a complex challenge. Toxicity assays are in majority conducted in mammals, due to their recognized efficacy at predicting toxicity in humans. However, mammalian assays are expensive, strictly regulated by law, and time-consuming. Given the morphological and developmental similarities among vertebrates, fish are a relevant test alternative to mammals [1,2]. For various reasons, the zebrafish embryo is particularly attractive in toxicology and pharmacology [3]. First, its transparency allows the visual detection of malformations, without interrupting development or invasive interventions [4]. Second, the number of eggs laid is high, and its development time is short [5]. Third, it is easy to maintain in the laboratory [6]. Finally, there are considerable gene homologies and neurophysiological similarities between zebrafish, mammals and humans [7,8]. Therefore, evidence obtained with zebrafish embryos should at least partially be translatable to humans [9]. From a regulatory point of view, until the age of 120 h, zebrafish embryos are an alternative to experiments with adult vertebrate species, because they are not protected by European animal welfare legislation until five days post-fertilization (hpf) [10,11].

Yet, extrapolation of toxicity from zebrafish to humans requires, at least, accounting for differences in pharmacokinetics between the two species [12]. Such differences might translate into differences in target organ concentrations for the same systemic exposure dose. In addition, knowing chemical concentrations in organs is fundamental to understand dose-response relationships [13,14]. Internal concentrations are often difficult to measure, but physiologically-based pharmacokinetic (PBPK) models can estimate them [15,16]. PBPK models connect anatomy, physiology, and biochemical processes to understand and compute a chemical's fate in the body. They allow approximate predictions of chemicals' concentration-time profiles in experimentally inaccessible organs from minimal data. Thereby, they provide mechanistic insight into toxicity and help reduce time, cost and need for animal experiments [17].

PBPK models have been published for adult zebrafish, mostly for ecotoxicological risk assessment [18–20]. Recently, Brox et al. [21] used a one-compartment two-parameter model developed by Gobas and Zhang [22] to explore the impact of physicochemical properties of polar compounds and of biological processes on embryo concentrations. This model, however, cannot describe decreases in concentrations due to metabolism [23] or dilution by organ growth, and Brox et al. concluded the necessity of more sophisticated toxicokinetic models for the zebrafish embryo.

In order to better explain, predict, and extrapolate developmental toxicity observed in zebrafish embryos, we developed a generic PBPK model integrating organ growth and hepatic metabolism. The model assumes quasi steady-state distribution between zebrafish cells, lysosomes, and mitochondria in different tissues (yolk, liver, gut, muscle, skeleton, eye, brain, heart, skin, and lumped other tissues). The model is generic in that it can simulate the distribution of many chemicals in zebrafish embryos on the basis of their physicochemical properties: chemicals' partition coefficients between cells or organelles and culture medium are calculated with the Simcyp® virtual in vitro intracellular distribution (VIVD) model [24]. The model can therefore be used for high-throughput predictions of internal concentrations in zebrafish.

We applied our model to the analysis of developmental toxicity data on valproic acid (VPA) and nine of its analogs: 2,2-dimethylvaleric acid, 2-ethylbutyric acid, 2-ethylhexanoic acid, 2-methylhexanoic acid, 2-methylpentanoic acid, 2-propylheptanoic acid, 4-eneVPA, 4-pentenoic acid and hexanoic acid. VPA is a notorious teratogenic antiepileptic and thymoregulator, inducing neural tube defects in mammalian embryos, probably by inhibition of histone deacetylase, interference with folate metabolism and inducing oxidative stress. However, the mechanism of action remains not well known [25–27]. VPA exerts its pharmacological effects mainly in the central nervous system by inhibition of the citric acid cycle and elevation of γ -aminobutyric acid (GABA) level [26]. Chemicals with similar structure can have similar properties, but for toxicological properties this should be backed-up by *in silico* and *in vitro* evidence. We demonstrate how the model can be used to base the toxicity ranking of VPA and the above analogs on internal concentration estimates rather than on nominal water medium concentrations, thereby helping transferability of zebrafish embryo test results to human risk assessment. In addition, we use the data published by Brox et al. [21] on 16 other chemicals to discuss the model performance for a larger class of chemicals.

2. Materials and methods

2.1. EU-ToxRisk zebrafish experiments

2.1.1. Test chemicals

Except for 4-eneVPA (Santa Cruz Biotechnology, Dallas, Texas, USA; 98 % purity), all other chemicals (valproic acid, 2,2-dimethylvaleric acid, 2-ethylbutyric acid, 2-ethylhexanoic acid, 2-methylhexanoic acid, 2-methylpentanoic acid, 2-propylheptanoic acid, 4-pentenoic acid and hexanoic acid) were purchased at the highest purity available from Sigma (Deisenhofen,

Germany). After initial range-finding tests, the ten chemicals were tested at three to eight different concentrations prepared from a 100 % DMSO stock solutions ($n = 3$). Nominal and analytically measured water concentrations as well as measured total embryo concentrations are summarized in Table S1 of the Supporting Information. The highest concentration tested led sometimes to 100% mortality of the embryos. In such cases, no concentration measurement was made.

2.1.2. Fish embryo toxicity testing

Adult wild-type zebrafish (*Danio rerio*) of the ‘Westaquarium strain’ were kept at the fish facilities of the Aquatic Ecology and Toxicology Group at the University of Heidelberg (licensed under no. 35-9185.64/BH). Based on OECD test guideline 236 as well as on complementary published work [28–31], embryos were raised and exposed until an age of 120 hours post-fertilization (hpf). According to the current EU animal welfare legislation, exposure of zebrafish may be extended to 120 hpf in cases of inconclusive observations until 96 hpf [11].

For initiation of the tests, embryos were immersed in the test solutions at the 16 cell-stage at the latest (≤ 90 min; before cleavage of blastodisc). To start exposures with minimum delay, twice the number of eggs eventually needed per treatment group were picked from the same batch of eggs and transferred into 100 ml crystallization dishes with the test concentrations or negative (artificial water according to ISO 7346-3) and positive controls (24.7 μM of 3,4-dichloroaniline) [33]. At 3 hpf at the latest, viable eggs were selected for normal development under the stereomicroscope (≥ 30 -fold magnification) and transferred to a final volume of 1 ml into 24 well plates (one embryo per well), which had been pre-exposed to the test solutions for 24 h to account for potential adsorption of the test solutions to the plastic walls of the wells. Test solutions were replaced at 24, 48, 72 and 96 hpf (without changing the well plates). Embryos were not dechorionated and hatched on their own at approximately 72 hpf.

Prior to replacement of the test solutions as well as at completion of the test (120 h), embryos were analyzed for macroscopically discernable alterations including the four morphological core endpoints listed by OECD TG 236 (coagulation, non-detachment of the tail, non-formation of somites and lack of heartbeat) [33] as well as any additional sublethal observation such as scoliosis/lordosis, eye deformation, loss of pigmentation, various types of edemata and general skeletal deformations [6,34,35].

For documentation, morphological alterations were recorded with a Zeiss Axio Cam ICc1 camera mounted on a Zeiss Olympus CKX41 microscope (Carl Zeiss, Oberkochen, Germany)

and analyzed using the Zeiss imaging program Zen lite 2011. After termination of the exposure, embryos were anesthetized in 2 ml Eppendorf cups immersed into crushed ice for 30 min, washed three times with artificial water to remove superficial chemicals from the embryo bodies and shock-frozen with liquid nitrogen for subsequent chemical analysis of internal doses.

2.1.3. Toxicity data analysis

Effective concentrations (EC) leading to 10, 20 and 50% of mortality or malformations (grouped together as “total effect”) were calculated using ToxRat Prof. Vers. 2.10 (ToxRat Solutions, Alsdorf, Germany).

2.1.4. Chemical analysis of actual chemicals’ concentration

Concentrations in the embryo and in the culture medium at 72 or 120 hpf for the ten VPA analogs studied were measured by liquid chromatography and mass spectrometry. 100 µL ultrapure water were added to each vial containing 2 to 10 zebrafish embryos and mixed thoroughly. To 100 µL of embryo, incubated water of standard curve sample, 10 µL of rosuvastatin (internal standard, IS, Sigma-Aldrich) in 50% methanol (VWR) were added, 300 µL or acetonitrile (Fisher Scientific) were then added to each sample. The samples in tube were centrifuged at 3000 rpm for 10 minutes, those in plates were centrifuged at 13,000 rpm for 5 minutes. 300 µL of supernatant were transferred to a fresh plate and subjected to a dry down under a stream of nitrogen at 50°C for approximately 30 minutes (until approximately 50 µL sample remained). To each well or sample, 100 µL ultrapure water were added, the plate was sealed and mixed thoroughly. The sample plates were then placed into the autosampler attached to a Sciex TripleTOF 6600 Quadrupole Time-Of-Flight (QTOF) mass analyzer (AB Sciex, Singapore). The conditions used when running the samples are shown in Table S2 of the Supporting Information.

2.2. Brox et al. zebrafish experiments

The 17 chemicals studied by Brox et al. [21] were 2,4-diclorophenoxyacetic acid, atropine, benzocaine, caffeine, chloramphenicol, cimetidine, clofibrilic acid, colchicine, cyclophosphamide, metoprolol, metribuzin, phenacetin, phenytoin, sulfamethoxazole, theophylline, thiacloprid and valproic acid. For each compound, Brox et al. performed two independent experiments, except for three compounds for which a third one was done to reduce experimental uncertainty. The experiments were performed on zebrafish embryos from 4 hpf (four-cell stage) to 96 hpf, exposed to external concentrations comprised between 10 to 250 mg/L, without medium change. The approximate stability of external concentrations during the experiments was checked. The quantity of chemical in each embryo was measured at 24, 48, 72 and 96 hpf. For metribuzin, phenacetin and benzocaine, measurements at 6 hpf were also performed. Three vials of 20 mL containing nine dechorionated embryos in 18 mL exposure solution, and three control vials, were used for each sampling time. Replicates were processed separately. The experiment was performed in a closed system at 26 ± 1 °C.

2.3. PBPK model structure

Figure 1 presents the schematic structure of our zebrafish embryo model. Ten compartments are considered because of interest of compound effect on their development: yolk, liver, skeleton, gut, eye, brain, heart, skin, muscles, other organs and tissues. Mitochondria and lysosomes in tissues (except the yolk) form two additional compartments. These organelles have a specific critical pH of 4.5 for lysosomes and 8 for mitochondria, leading to potential ‘ion trapping’ phenomena (sequestration of compounds because of their differential ionization between organelles and cellular embryo). The model also considers air and plastic to medium partition. Protein binding and ionization in the medium are also accounted for. In the experiments reported in this article the culture medium contained no proteins; therefore, protein binding was turned off in the model. Instantaneous diffusion across the various compartments is assumed. Metabolic clearance is modeled as a dynamic process in the liver. Organ volume growth with time is also modeled. Therefore, chemical’s concentrations and quantities in the model change with time, even though (for notation simplicity) time indexing is not always explicit in the following equations.

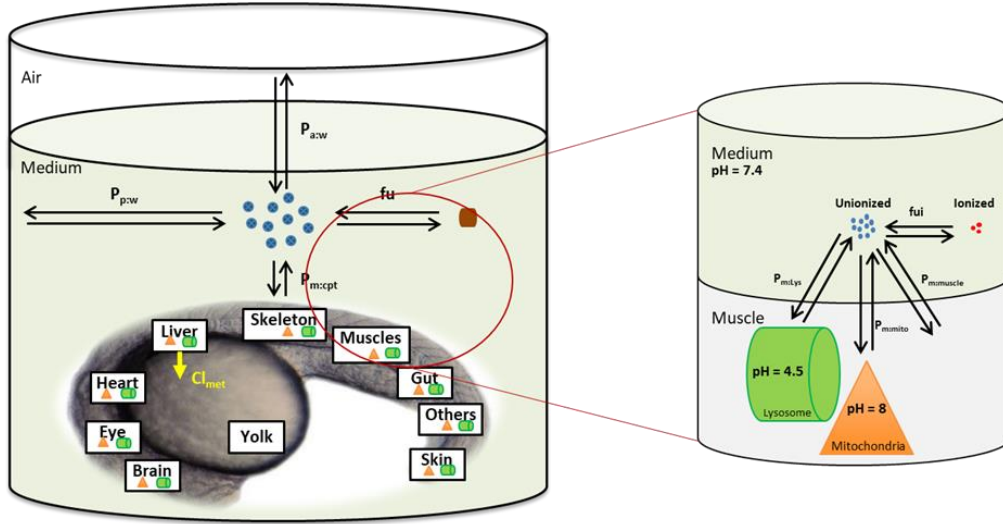


Figure 1. Structure of the zebrafish embryo model. The chemical of interest partitions between the various compartments and can be metabolized in the liver. It can also partition to the air and bind to the plastic walls and culture medium proteins. Legend: Cl_{met} : Metabolic clearance; $P_{p:w}$: Plastic to water partition coefficient; $P_{a:w}$: Air to water partition coefficient; $P_{m:cpt}$: Compartment to medium partition coefficient; f_u : Fraction unbound in medium; f_{ui} : Fraction unionized in medium.

Concentrations in organs, yolk, lysosomes and mitochondria (C_i), are assumed to be at any time proportional to the concentration unbound in medium ($C_{medium,u}$). The proportionality factors are the organs, yolk, lysosomes or mitochondria over medium unbound partition coefficients ($P_{i:mu}$) eventually corrected by a scaling factor (f_{pc}):

$$C_i = C_{medium,u} \times f_{pc} \times P_{i:mu} \quad (1)$$

$C_{medium,u}$ is computed according to:

$$C_{medium,u} = \frac{Q_{parent}}{\frac{V_{medium}}{f_u} + P_{a:w} \times f_{ui} \times V_{air}(t) + P_{p:w} \times S_{medium}(t) \sum (P_{j:mu} \times V_j(t))} \quad (2)$$

where V_{medium} is the volume of culture medium, f_u the fraction unbound in medium, f_{ui} the fraction unionized in medium, $P_{a:w}$ and $P_{p:w}$ are respectively the air and plastic to water partition coefficient, $V_{air}(t)$ the volume of air in head space at time t , $S_{medium}(t)$ the surface area of medium in contact with plastic. $V_j(t)$ are the volumes of yolk, liver, gut, muscle, skeleton, eye, brain, heart, skin, and other tissues at time t . The total concentration in medium (C_{medium}) is:

$$C_{medium} = \frac{C_{medium,u}}{f_u} \quad (3)$$

Since we model the developing embryo, it is necessary to consider organ growth. The embryo volume without yolk at time t , $V_{embryo}(t)$, is computed as:

$$V_{embryo}(t) = V_{embryo}(120) \times \sum sc_k(t) + V_{embryo}(0) \quad (4)$$

where $V_{embryo}(120)$ is the volume of embryo at 120 hpf, $sc_k(t)$ represents the fraction of $V_{embryo}(120\text{hpf})$ taken up by organ k (liver, gut, muscle, skeleton, eye, brain, heart, skin, or other tissues) at time t , and $V_{embryo}(0)$ is the embryo volume at the start of the experiment.

$V_{embryo}(0)$ is computed by assuming that the fertilized egg is a half-sphere of radius $r_{embryo,0}$ equal to 0.13 mm [36]:

$$V_{embryo}(0) = \frac{2}{3} \times \pi r_{embryo,0}^3 \quad (5)$$

The yolk volume at time t , $V_{yolk}(t)$, is given by:

$$V_{yolk}(t) = V_{yolk}(0) \times \exp(-K_{d,yolk} \times t) \quad (6)$$

where $K_{d,yolk}$ is the yolk consumption rate constant, and $V_{yolk}(0)$ is the yolk volume at start of experiment, computed according to the hypothesis that it is a sphere of radius $r_{yolk,0}$ equal to 0.4 mm [36].

$$V_{yolk}(0) = \frac{4}{3} \times \pi r_{yolk,0}^3 \quad (7)$$

The total volume of the embryo, $V_{embryo\text{total}}(t)$, is the sum of $V_{yolk}(t)$ and $V_{embryo}(t)$.

$$V_{embryo\text{total}}(t) = V_{yolk}(t) + V_{embryo}(t) \quad (8)$$

The organ volumes at time t , $V_k(t)$, for liver, gut, muscle, skeleton, eye, brain, heart, skin, and other tissues are computed as:

$$V_k(t) = V_{embryo}(120) \times sc_k(t) \quad (9)$$

The fractional volumes $sc_k(t)$ were computed for each organ according to the time-dependent equations:

$$sc_k(t) = \exp(K_{g,k} \times (t - \tau_k)) - 1 \quad (10)$$

where τ_k is the time of growth initiation for organ k . Before τ_k , organ k volume is null. The organ growth rates ($K_{g,k}$) were calibrated using published information and our own data on embryos (see next section).

The surface area of medium in contact with plastic, $S_{\text{medium}}(t)$, is:

$$S_{\text{medium}}(t) = 4 \times \frac{V_{\text{content}}(t)}{D_{\text{well}}} + \pi \left(\frac{D_{\text{well}}}{2} \right)^2 \quad (11)$$

where D_{well} is the well diameter, and $V_{\text{content}}(t)$ is the total volume of content per well, computed as:

$$V_{\text{content}}(t) = V_{\text{medium}} + V_{\text{embryo}_{\text{total}}}(t) \quad (12)$$

$V_{\text{air}}(t)$ is computed as the difference between well volume and medium, embryo and yolk volumes:

$$V_{\text{air}}(t) = V_{\text{well}} - V_{\text{content}}(t) \quad (13)$$

The air concentration depends on the air to water partition coefficient, eventually corrected by the scaling factor (f_{pc}), and on the fraction unbound and unionized in medium:

$$C_{\text{air}} = f_{\text{pc}} \times P_{\text{a:w}} \times C_{\text{medium,u}} \times f_{\text{u}_u} \quad (14)$$

Similarly, the quantity of chemical bound to the culture walls per unit surface area depends on the plastic to water partition coefficient (which has the dimension of a length):

$$C_{\text{plastic}} = P_{\text{p:w}} \times C_{\text{medium,u}} \quad (15)$$

If metabolism is assumed to be linear in the embryo, the total quantity of metabolites formed per unit time in system is proportional to the number of liver cells of the embryo (N_{cell}), metabolic clearance per liver cell (Cl_{met}) and parent chemical concentration in liver cells (C_{liver}):

$$\frac{Q_{\text{met}}}{dt} = N_{\text{cells}} \times Cl_{\text{met}} \times C_{\text{liver}} \quad (16)$$

If metabolism is assumed to be saturable, a Michaelis-Menten term is used and Eq. 16 becomes:

$$\frac{Q_{\text{met}}}{dt} = N_{\text{cells}} \times V_{\text{max}} \times C_{\text{liver}} / (K_m + C_{\text{liver}}) \quad (17)$$

where V_{max} is the maximum rate of metabolism and K_m the Michaelis-Menten constant.

The total quantity of parent molecules in the system (Q_{parent}) depends of quantities in medium, air, total embryo and plastic. The model can account for the saturation of plastic binding by pre-incubation with the test substance prior to embryo exposure. The model splits Q_{parent} into two quantities, both function of time. A labile quantity present in medium and air (Q_{labile}), reset at 0

at each medium change; and a fixed quantity in embryo and bound on plastic walls (Q_{fixed}), impervious to medium renewals.

$$Q_{parent} = Q_{labile} + Q_{fixed} \quad (18)$$

$$\frac{Q_{labile}}{dt} = -f_{labile} \times \frac{Q_{met}}{dt} \quad (19)$$

$$\frac{Q_{fixed}}{dt} = -(1 - f_{labile}) \times \frac{Q_{met}}{dt} \quad (20)$$

where f_{labile} is the fraction of Q_{parent} in medium and air:

$$f_{labile} = \frac{\frac{C_{medium,u}}{f_u} \times V_{medium} + C_{air} \times V_{air}(t)}{Q_{parent}} \quad (21)$$

N_{cell} changes with time and is calculated from liver volume at time t , $V_{liver}(t)$, and hepatocyte volume (V_{hep}):

$$N_{cells} = \frac{V_{liver}(t)}{V_{hep}} \quad (22)$$

The quantity in medium (Q_{medium}) depends on concentration unbound in medium, adjusted by the fraction unbound:

$$Q_{medium} = \frac{C_{medium,u} \times V_{medium}}{f_{u_u}} \quad (23)$$

The quantities in organs, yolk and air (Q_i) are computed as:

$$Q_i = C_i \times V_i(t) \quad (24)$$

The quantity in embryo, excepting yolk, Q_{embryo} , is the sum of organ quantities, and the concentration in embryo (C_{embryo}) is computed as:

$$Q_{embryo} = \sum Q_{organ} \quad (25)$$

$$C_{embryo} = \frac{Q_{embryo}}{V_{embryo}} \quad (26)$$

The total quantity of compound in the embryo, $Q_{embryo,total}$, is the sum of $Q_{yolk}(t)$ and Q_{embryo} :

$$Q_{embryo_{total}} = Q_{yolk} + Q_{embryo} \quad (27)$$

The quantity bound to plastic ($Q_{plastic}$) is given by:

$$Q_{plastic} = C_{plastic} \times S_{medium} \quad (28)$$

The quantity in lysosomes (Q_{lyso}) and mitochondria (Q_{mito}) depend on the volume of the embryo and the fractions of lysosome (f_{lyso}) and mitochondria (f_{mito}) in cells, respectively:

$$Q_{lyso} = C_{lyso} \times V_{embryo}(t) \times f_{lyso} \quad (29)$$

$$Q_{mito} = C_{mito} \times V_{embryo}(t) \times f_{mito} \quad (30)$$

The length of the embryo was calibrated by ourselves with data from Kimmel et al. [36], using the following empirical equation:

$$L_{embryo} = A \times \frac{t^B}{C^B + t^B} + D \quad (31)$$

where $A = 0.0260$ dm, $B = 4.397$, $C = 1617$ minutes and $D = 0.00755$ dm.

At the start of a simulation, time zero, the quantity of parent compound (Q_{parent}) in the system is known. It includes what has been introduced in the medium plus what is already attached to the plastic walls in case of pre-incubation (Eq. 15 applies then). Therefore, the fraction f_{labile} can be computed (Eq. 21) and used to set the initial values for the state variable Q_{labile} and Q_{fixed} , as corresponding fractions of Q_{parent} .

Medium renewals introduce a discontinuity in the solution of the two differential equations of the model (for Q_{labile} and Q_{fixed}). That case is handled classically: Integration is stopped just at the renewal time; Q_{fixed} is kept at its current value and Q_{labile} is made equal to the new quantity introduced in the system; Simulation is restarted from the renewal time.

2.4. PBPK model parameters

2.4.1. Partition coefficients

To estimate the fraction of chemical unbound in medium (f_u), the fraction unionized in medium (f_{ui}), the plastic to water partition coefficient ($P_{p:w}$), the air to water partition coefficient ($P_{a:w}$), and the organs (liver, gut, muscle, skeleton, eye, brain, heart, skin, others, yolk, lysosomes and mitochondria) to medium unbound partition coefficient ($P_{mu:i}$), we use the Simcyp® VIVD model [24]. It computes parameter values on the basis of physicochemical properties of the substance considered ($\log P$; Henry's constant; pK_a ; compound's character: mono or dibasic, mono or diacid, neutral, ampholyte; molecular weight; blood to plasma ratio and fraction unbound in bovine serum, pH and membrane potential) and embryo organ properties, obtained from the literature on embryos (for the yolk) [37,38] or adult fish [18]. As

there was no bovine serum or other proteins in the zebrafish culture medium, f_u was set equal to 1 for all compounds.

2.4.2. Physiological parameters

The model's physiological parameters are given in Table S3. Organ growth rates were estimated from our own data on total embryo volume and volume without yolk at different times (see Figure S1 in Supporting Information) using the following procedure: The yolk consumption rate constant was estimated using a simple exponential decay equation. The embryo's volume at 120 hpf ($V_{embryo(120hpf)}$), the fractional volume of muscle at 120 hpf ($sc_{muscle(120hpf)}$) and the "other organs" fractional volume at 120 hpf ($sc_{others(120hpf)}$) were estimated by fitting the organ growth part of the model to the data. Calibration was performed by MCMC simulations in a Bayesian statistical framework [39–41]. The data was assumed to be log-normally distributed around the model predictions with a geometric standard deviation σ (estimate of residual uncertainty). Non-informative uniform priors were used for the three parameters to calibrate, so as to "let the data speak". Two MCMC chains of 10000 iterations were simulated and one of every two random samples produced were recorded. Convergence of the two chains was assessed using Gelman and Rubin's Rhat convergence criterion [42].

The fractional volumes, $sc_k(120hpf)$, for the rest of the organs were then computed by rescaling their literature values [18,43]:

$$sc_k(120hpf) = sc_{k\text{ literature}}(120hpf) \times \frac{(1 - sc_{muscle\text{ fitted}}(120hpf) - sc_{others\text{ fitted}}(120hpf))}{(1 - sc_{muscle\text{ literature}}(120hpf) - sc_{others\text{ literature}}(120hpf))} \quad (32)$$

Organ growth rates were finally obtained by:

$$K_{g,k} = \frac{\ln(sc_k(120hpf) + 1)}{(t_{final} - \tau_k)} \quad (33)$$

t_{final} being equal to 120 hours.

2.4.3. Estimation of metabolic clearance and partition coefficient scaling factor

To improve the model fit to the data beyond that obtained with *ab initio* predictions, metabolic clearance (Cl_{met}) and the scaling factor (f_{pc}) were calibrated for individual chemicals on the basis of chemical concentration data in the embryo and medium. Fits were performed with data obtained at 120 hpf (and at 72 hpf for VPA). Those parameters were calibrated jointly, using MCMC simulations. The data on C_{medium} and C_{embryo} was assumed to be log-normally

distributed around the model predictions (taken as geometric mean) with geometric standard deviation σ . The two non-detectable concentration data points were excluded from the analysis. The SD σ was also calibrated by sampling and assumed to be *a priori* distributed normally around 1.5 ± 1.5 SD with a truncation from 1.5 to 10 (that is, between 50% error and a 10-fold error at most). Two Markov chains of 10000 iterations were simulated for each chemical, and one in two random parameter samples were recorded. The last half of each recorded set of samples was kept and convergence of the two chains was assessed using Gelman and Rubin's convergence criterion [42]. The prior distribution of metabolic clearance was assumed to be uniform (*i.e.*, uninformative) from 0 to either 10^{-11} , 10^{-10} , or 10^{-9} L/min, depending on compound (see Table 1). Different upper bounds were used to speedup convergence of the chains toward the posterior distribution, but they do not affect the posterior estimates as they were not reached during sampling at convergence. The prior distribution of f_{pc} was assumed to be uniform from 0 to 5 (*i.e.*, uninformative).

2.5. Software

The static model equations of the VIVD model were coded in R version 3.4.3 [44]. The corresponding R script was used as a preprocessor to obtain chemical-specific parameter values for input to the embryo model. All the dynamic model simulations and MCMC calibrations were performed with GNU MCSIM version 5.6.6 (<https://www.gnu.org/software/mcsim/>) [45]. The model code is given as supplemental material and on the web at <https://sites.google.com/site/modelecotoxtox/Software>.

3. Results and discussion

3.1. Physiological parameters' calibration

The model accounts for the embryo's organ growth over time. We estimated the yolk consumption rate constant, the embryo's volume at 120 hpf, the fractional volume of muscle at 120 hpf and the "other organs" fractional volume at 120 hpf on the basis of our experimental embryo volume data. Figure S1 (in Supporting Information) shows the observed and predicted organ growth over time. Predicted total embryo volume and embryo volume without yolk fit the data rather well: the median relative error for the embryo volume estimate without yolk is equal to 1.09, and for the total embryo volume it is equal to 0.94. The Figure S2 also shows the contribution of each organ to the total embryo volume. The estimated parameter values are given in Table S3 of the Supporting Information.

3.2. *Ab initio* predictions of embryo concentrations

The parameter values obtained with the VIVD model for VPA and its analogs are given in Table S4 (Supporting Information). The primary physicochemical properties input to the VIVD model for calculation of those values are given in Supplemental Table S5 (Supporting Information). 2-propyl-heptanoic acid had the highest octanol over water partition coefficient of all the analogs studied here. All analogs were predicted to be mainly present in ionized form in the medium. The partition coefficient for plastic binding is difficult to interpret directly because of its dimension (a length). Yet, it is useful to assess the impact of the materials used on the kinetics of the test chemicals. This is best done by comparing the predictions of the quantities bound to plastic and present in water. In our case, for all substances, the quantity bound to plastic was predicted to be about 0.2% of what is in water and therefore negligible.

The compounds were predicted to partition preferentially to the yolk, except for VPA and 2-propylheptanoic acid, which partition to water rather than yolk. They also all have higher affinity for the other embryo tissues than for medium. Affinity for lysosomes was predicted to be ten to a hundred times higher than for the other compartments for all compounds, except again for 2-propylheptanoic acid.

Figure S2 (Supporting Information) shows the difference of organ concentrations as a function of time for VPA. We can see that VPA concentrates preferably in mitochondria and weakly in lysosomes.

We compared the observed analogs' concentrations to those obtained when using VIVD-predicted parameters. For that case, metabolic clearance (Cl_{met}) was set to 0 L/min, because we did not have a way to predict it) and f_{pc} to 1. Figure 2 shows that those *ab initio* predictions tend to be underestimates of embryo concentrations (median relative error 0.11). The medium concentration data were well predicted (median relative error 0.90).

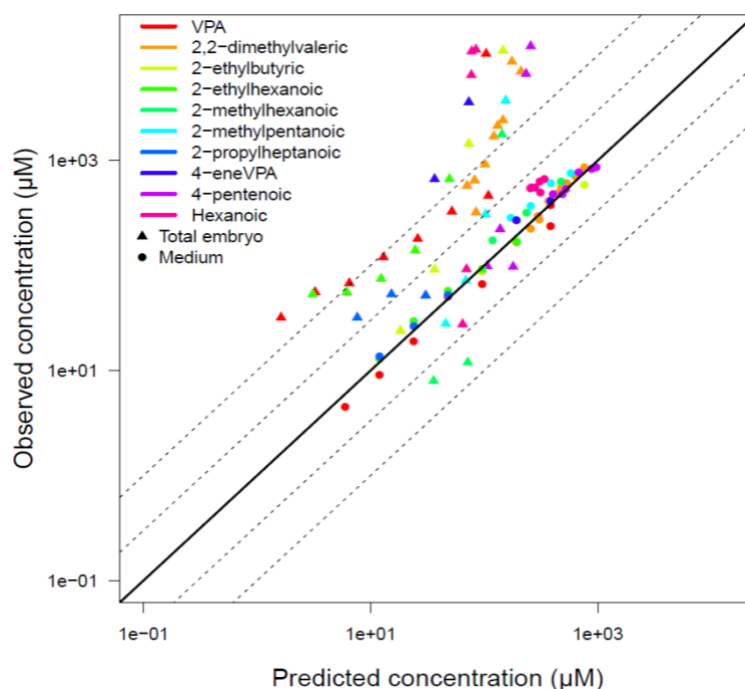


Figure 2. Observed *versus* model-predicted concentrations in the zebrafish embryo and in culture medium for valproic acid and nine analogs. The VIVD-computed parameter values were used without further adjustment. Metabolic clearance was set to zero and the correction factor of partition coefficient to one. The black line corresponds to perfect predictions. Dashed lines delineate the three- and ten-fold error bounds.

The VIVD model estimates of partition coefficients are certainly not perfect, for example due to imprecision affecting some input parameters, such as Henry's law constant. For low volatility compounds, the assumption of instantaneous partitioning between culture medium and air in the head-space may also over-predict distribution to the head-space. Conversely, particularly for high volatility compounds, the model will likely under-predict loss to the head-space if the experimental system is not hermetically sealed (as assumed by the model). Modelling distribution into the head-space air as a dynamic process, as we did for metabolism, might improve predictions [46]. However, that would complicate the model and add many parameters. The physicochemical tissue properties values (except for the yolk) were obtained in adult fish because we did not have embryo fish specific data: This should be improved with specific measurements. It should also be noted that the VIVD model has been developed for relatively well-behaved small molecules and does not have a universal domain of applicability. It does not predict metabolic clearance either. Actually, there are no published QSAR or other *in silico* methods to predict metabolic clearances in zebrafish embryo. QSAR methods have been developed to predict the biotransformation half-lives or rates in adult fish [47,48].

However, since there are important physiological differences between adult fish and embryos, embryos fall out of the applicability domain of these QSARs. It would be interesting to understand whether these QSARs nevertheless provide useful upper bounds or lower bounds on the embryo biotransformation parameters.

Note also that the VIVD model is quite general and considers a number of parameters that may be irrelevant for the modelling of zebrafish embryo internal concentrations in specific cases. For instance, under normal conditions, there should be no proteins in the exposure medium. We chose to keep those features for compatibility with the VIVD model and greater generality of our embryo model, so that it can be used for *ab initio* predictions for a large number of chemicals.

3.3. Improving predictions by fitting VPA and analogs EU-ToxRisk data

For the *ab initio* predictions, metabolism was neglected. This may be acceptable for VPA analogs, since neglecting metabolism should lead to overestimated internal concentrations, while in fact the predictions were underestimates. To check the assumption of negligible metabolism, we estimated Cl_{met} by calibration with the data. Observed *versus* best fit (maximum posterior probability) concentrations are presented on Supporting Information's Figure S3. Cl_{met} value. Predictions of total embryo concentrations were out the ten-fold interval error and under-estimated, while predictions of medium concentration were included in the ten-fold error interval.

Supplemental Figure S5 shows the observed and predicted concentrations of VPA and its analogs in medium and in the total embryo (including yolk) as a function of time, after estimating Cl_{met} only. Table S6 (Supporting Information) summarizes the posteriors distributions of Cl_{met} and σ . The maximum posterior estimates are the most likely, best fitting, values. For the various compounds, the metabolic clearance best estimates were of the order of 10^{-15} to 10^{-13} L/min, which for an embryo volume of about 3×10^{-7} L correspond to half-lives well above 3500 hours (about 150 days). This implies negligible metabolism of VPA and analogs in our zebrafish embryos and validates our *ab initio* choice of null metabolic clearances.

A better way to improve the internal concentration estimates of VPA analogs was to adjust the VIVD predicted partition coefficients using the concentration measurements. MCMC simulations were used to calibrate Cl_{met} and f_{pc} together on the basis of the data. Figure 3 plots observed *versus* predicted concentrations in that case. Here, the model predictions were also obtained using the best fitting parameter values. Overall, the points were better aligned with the

perfect fit line and mostly comprised within the ten-fold error interval. Hexanoic acid, 4-pentenoic, 2-methylhexanoic acid and 2-methylpentanoic acid were the four worst predicted chemicals and did not fall in the three-fold error interval. The medium concentrations were under-predicted by the model with a median relative error of 0.72. The embryo concentrations were also slightly under-predicted with a median relative error of 0.83. The best estimate of σ corresponded on average to a factor 2.5.

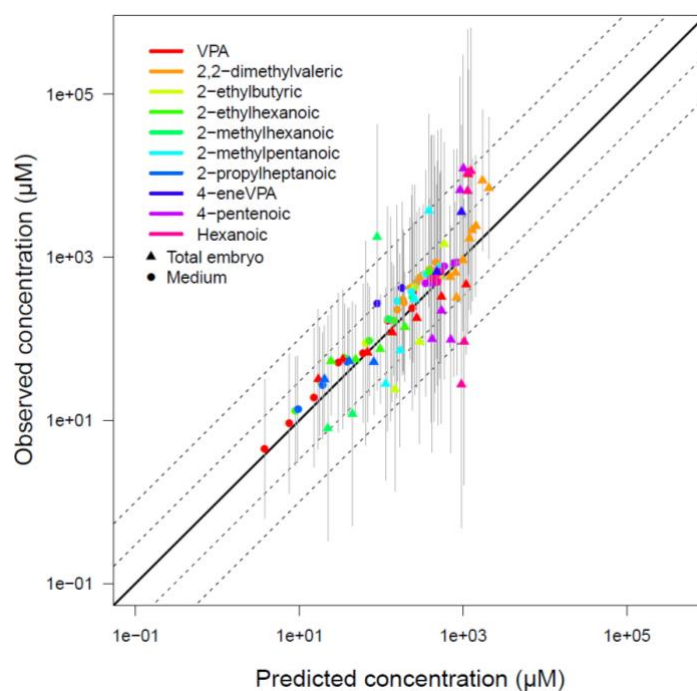


Figure 3. Observed *versus* predicted concentrations of valproic acid (VPA) and nine analogs, in the zebrafish embryo and culture medium. This is the best fit obtained after Cl_{met} and f_{pc} were simultaneously estimated. The solid black line corresponds to perfect fit. Dashed lines correspond to the three- and ten-fold error intervals. The grey bars correspond to $\pm$ one residual SD (σ).

Figure 4 presents the observed and predicted concentrations of VPA in medium and in the total embryo (including yolk) as a function of time, following exposure to various nominal concentrations, after simultaneous estimation of Cl_{met} and f_{pc} . Since the model was linear with dose, all concentrations were normalized to a nominal concentration of 1 mM to simplify the Figures (one model estimate only is needed for all doses). The discontinuities of the concentration-time curve are due to the daily changes of culture medium. Similar results are shown for the VPA analogs in Supplemental Figure S4. Almost all predicted medium and embryo time-course concentrations are in the 95% confidence interval of the model predictions

(but those intervals can be large, showing a large uncertainty in measurements and consequently in model predictions).

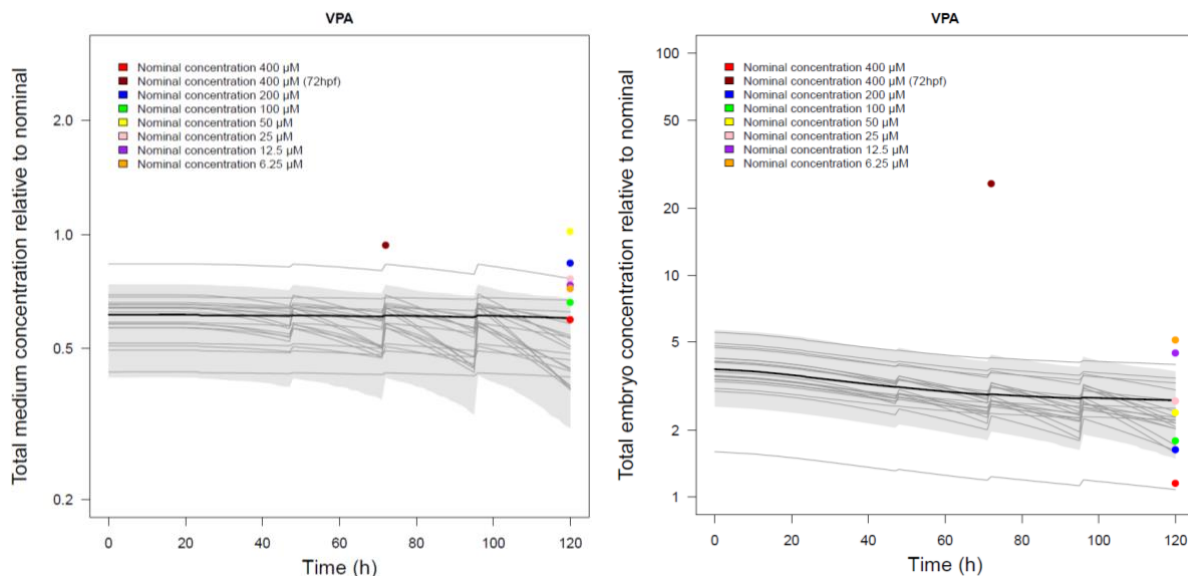


Figure 4. Predicted (lines) and observed (dots) VPA concentrations in medium (left) and in total embryo (right) as a function of time, after estimating Cl_{met} and f_{pc} . All concentrations were normalized to the nominal concentration for plotting. Culture medium was replaced every day. The grey area defines the 95% confidence interval. The thick black line is the maximum posterior predicted concentration time-course. The thin lines are 20 predictions obtained using random parameter vectors drawn from their posterior distribution.

Table 1 summarizes, for each chemical, the posterior distributions of Cl_{met} , f_{pc} and uncertainty SD σ obtained by MCMC calibration with the concentration data. For the various compounds, the metabolic clearance best estimates were between 10^{-12} to 10^{-16} L/min. The highest upper confidence limit on a clearance was found for 2-Methylhexanoic acid, which for an embryo volume of about 3×10^{-7} L corresponds to a half-life of about 11 hours. The half-lives corresponding to the best estimates would be considerably higher. Therefore, the conclusion of negligible metabolism for those compounds in the zebrafish embryo appears to be coherent. The best estimates of the partition coefficient scaling factor, f_{pc} , are in the range of 0.608 (for 2-methylhexanoic acid) to 27.9 for 4-eneVPA. According to their 95% confidence intervals, the f_{pc} values for all compounds are significantly different from 1. It appears that the medium over tissue partition coefficient predicted by the VIVD model had to be increased to improve data fit for the compounds of interest. The values of σ , corresponding on average to a factor 2.5, shows large uncertainties in measurements and modeling. 2-methylhexanoic acid, 4-pentenoic acid and hexanoic acid are particularly affected, with uncertainties higher than a factor 3.

Causes for such large uncertainties are multiple and cumulate their effects: the precision of the measurement method is limited, there is variability between embryos, initial concentration in the medium may be different from the nominal concentration, there may be some loss of the substance in the air, loss by degradation other than metabolic, loss or amplification during sample preparation, *etc.*

Table 1. Estimation of means, standard deviation (SD), 95% confidence intervals (IC95%) and maximum posterior (MP) value for the metabolic clearance, the partition coefficient correction factor, and residual uncertainty SD σ , for VPA and its analogs.

Compound	CL_{met} (L/min)			f_{pc}			σ		
	Mean $\pm$ SD	IC 95%	MP	Mean $\pm$ SD	IC 95%	MP	Mean $\pm$ SD	IC 95%	MP
Valproic acid**	$3.55 \times 10^{-12} \pm 3.10 \times 10^{-12}$	$[1.38 \times 10^{-13}; 1.14 \times 10^{-11}]$	3.49×10^{-13}	20.0 ± 7.04	$[9.46; 36.9]$	16.8	2.27 ± 0.397	$[1.74; 3.26]$	1.97
2,2- Dimethylvaleric acid***	$1.67 \times 10^{-12} \pm 1.81 \times 10^{-12}$	$[5.25 \times 10^{-14}; 7.08 \times 10^{-12}]$	8.13×10^{-16}	2.28 ± 0.377	$[8.61; 30.4]$	15.9	2.28 ± 0.377	$[1.77; 3.21]$	2.00
2-Ethylbutyric acid**	$1.23 \times 10^{-11} \pm 1.20 \times 10^{-11}$	$[3.29 \times 10^{-13}; 4.32 \times 10^{-11}]$	5.59×10^{-13}	16.7 ± 10.4	$[3.43; 43.5]$	11.9	3.44 ± 0.739	$[2.32; 5.24]$	2.89
2-Ethylhexanoic acid***	$2.49 \times 10^{-12} \pm 2.20 \times 10^{-12}$	$[6.32 \times 10^{-14}; 8.36 \times 10^{-12}]$	4.16×10^{-14}	11.7 ± 3.24	$[6.19; 18.9]$	10.6	1.87 ± 3.24	$[1.52; 2.73]$	1.54
2-Methylhexanoic acid*	$8.67 \times 10^{-11} \pm 8.89 \times 10^{-11}$	$[1.80 \times 10^{-12}; 3.17 \times 10^{-10}]$	1.34×10^{-12}	1.76 ± 1.66	$[0.257; 6.86]$	0.608	3.79 ± 0.803	$[2.58; 5.70]$	3.18
2-Methylpentanoic acid**	$2.06 \times 10^{-11} \pm 1.71 \times 10^{-11}$	$[8.42 \times 10^{-13}; 6.65 \times 10^{-11}]$	1.76×10^{-12}	3.90 ± 2.12	$[1.04 ; 8.96]$	2.71	3.23 ± 0.693	$[2.22; 4.91]$	2.71
2-Propylheptanoic acid***	$3.67 \times 10^{-12} \pm 2.62 \times 10^{-12}$	$[1.24 \times 10^{-13}; 9.41 \times 10^{-12}]$	3.43×10^{-13}	4.17 ± 1.78	$[1.67; 8.64]$	3.13	2.02 ± 0.486	$[1.51; 3.33]$	1.51
4-eneVPA***	$8.15 \times 10^{-12} \pm 8.15 \times 10^{-12}$	$[1.70 \times 10^{-13}; 3.09 \times 10^{-11}]$	2.67×10^{-14}	35.4 ± 22.7	$[6.29; 92.5]$	27.9	3.33 ± 0.799	$[2.07; 5.11]$	2.43
4-Pentenoic acid**	$1.51 \times 10^{-11} \pm 1.41 \times 10^{-11}$	$[6.73 \times 10^{-13}; 5.12 \times 10^{-11}]$	3.09×10^{-12}	6.93 ± 4.10	$[1.60; 17.5]$	4.55	3.59 ± 0.669	$[16.5; 19.2]$	3.20
Hexanoic acid***	$9.23 \times 10^{-12} \pm 8.65 \times 10^{-12}$	$[2.94 \times 10^{-13}; 3.46 \times 10^{-11}]$	3.59×10^{-13}	16.1 ± 11.3	$[3.05; 47.6]$	9.10	4.36 ± 0.769	$[3.15; 6.19]$	4.04

* Prior on CL_{met} : Uniform (0, 10^{-9}); prior on f_{pc} : Uniform (0, 5).

** Prior on CL_{met} : Uniform (0, 10^{-10}); prior on f_{pc} : Uniform (0, 5).

*** Prior on CL_{met} : Uniform (0, 10^{-11}); prior on f_{pc} : Uniform (0, 5).

A limitation of our model is that it does not consider the chorion nor active transport: The diffusion of chemicals within the embryo is assumed to be instantaneous. This assumption is reasonable given the small size of the embryo, but we do not yet have data to test its validity. We are currently working on improving the model to account for rate-limiting diffusions. Note that dechoriation could remove a significant fraction of chemical bound to the chorion. This would not affect the concentration measured in the embryo, but would prevent an assessment of mass balance and product loss (for example by metabolism) in the experiment if the chorionic concentration were not measured. It would be interesting to measure concentrations with and without the chorion to better understand its effects on the pharmacokinetics in the embryo. For transporters, there appear to be similarities between mammals and zebrafish embryo efflux transporters [49]. This may imply that active transport can play a role in modulating zebrafish embryos' exposure to xenobiotics, but we do not have sufficient data to verify it for the VPA analogs investigated here. Likewise, our modeling of organ growth was based on limited data on organogenesis and embryo volumes. There is room for improvement with specific measurements of embryo volumes as a function of time. Note also that the Monte-Carlo simulated confidence intervals we estimated and presented on the Figures only reflect parametric uncertainty, and not structural model uncertainty.

Figure S6 shows the observed concentrations of VPA and analogs in the total embryo as a function of nominal concentration. Despite the small number of data points to infer on saturation, we also considered Michaelis-Menten kinetics as an alternative to linear ones, except for VPA, 2-propylheptanoic acid, and 4-eneVPA, for which linearity seemed to apply. In the case of saturable metabolism, parameters V_{max} and K_m were calibrated together with f_{pc} .

Figure S7 presents the kinetic profile for 2-ethylbutyric acid when saturable metabolism is assumed. The 95% confidence intervals were large and data fits for medium and total embryo were not improved compared to linear kinetics. Metabolism remained negligible.

3.4. Prediction of the VPA data of Brox et al.

Without any additional adjustment, we performed simulations of Brox *et al.* data on VPA with our best parameter estimates for Cl_{met} and f_{pc} (Table 1). Figure 5 shows an over-prediction with a median relative error of a factor 2. A difference of this order is expected because our above parameterization is quite uncertain and the laboratories which produced the data used somewhat different methods (dechoriation in Brox *et al.* experiments, different analytical methods *etc.*) This is a limited validation of the model, but at least for VPA, it predicts

reasonably well a very different data set. Actually, some of the random prediction curves shown on Figure 5 are closer to the data. We examined the Cl_{met} and f_{pc} values leading to the four prediction curves closest to the data. Their means were 7.9 ± 2.6 picoL/min, and 13 ± 6 , respectively, while the means in Table 1 are 3.55 picoL/min for Cl_{met} , and 20 for f_{pc} . So, Brox data point to a somewhat higher metabolism of VPA and lesser correction of the VIVD partition coefficient estimates than the EU-ToxRisk data.

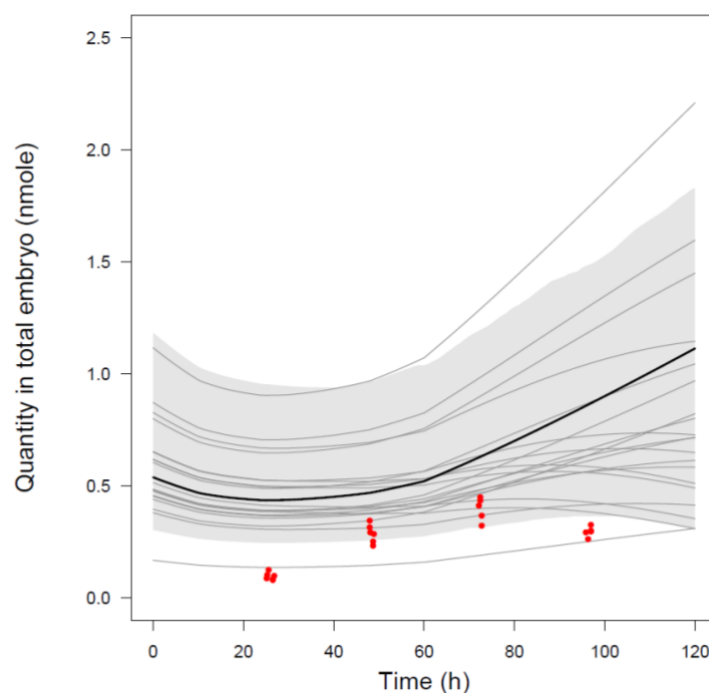


Figure 5. Predicted (lines) and data by Brox *et al.* (dots). VPA quantities in total embryo as a function of time, after estimating Cl_{met} and f_{pc} from our data. The culture medium was not replaced. The grey area defines the 95% confidence interval. The thick black line is the maximum posterior prediction. The thin lines are 20 predictions obtained using random parameter vectors drawn from their posterior distribution.

3.5. Fit of the model to Brox *et al.* data

For all chemicals assayed by Brox *et al.* we also calibrated the model. Figure 6 shows the observed *versus* predicted quantities per embryo. Model predictions were obtained using the best fitting parameter values. Most points fall within the three-fold error interval, showing that the model can describe the data reasonably well. All the kinetic profiles (Figure S8) and fitted parameter values (Table S7) are shown in Supplementary information. The model captures most of the time courses correctly for those compounds. There are some misfits (Metribuzin, Phenytoin, Thiocloprid), which could be due to chemicals actively transported, or not

penetrating the chorion, *etc.*, but additional model complexity or fitting would be needed to improve this. Note that the metabolic clearance values obtained (Table S7) are again very low, and the ups and downs of the concentration time-courses are sufficiently explained by changes in organ sizes and yolk consumption. It would be very interesting to confirm this poor metabolic capacity of the embryo by measuring the expected metabolites, but this is challenging, because minute amounts of metabolites are expected to be formed. It might be worth designing experiments, in which the volume of the water medium would be very small (and the experiment time relatively short) to avoid low concentrations of the soluble metabolites in the medium.

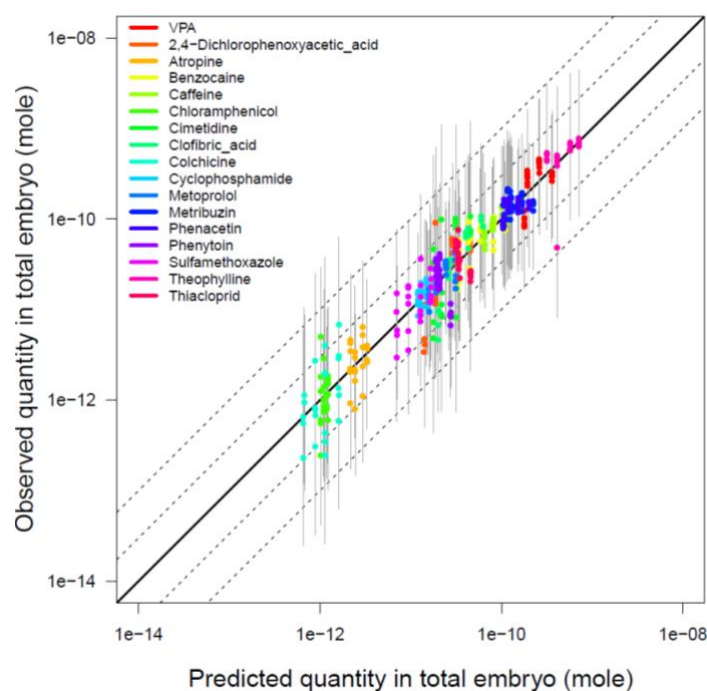


Figure 6. Observed *versus* predicted quantities in the embryo, for the 17 Brox *et al.* studied compounds. This is the best fit obtained with Cl_{met} , f_{pc} and $P_{a:w}$ simultaneous estimated. The black line corresponds to perfect fit. Dashed lines describe the three- and ten-fold error. The grey bars are the uncertainty σ .

3.6. Using the model to correct effective concentrations for pharmacokinetics

The main aim of our modeling effort was to estimate internal embryo concentrations in order to base effective concentration (EC) estimates on them – rather than on nominal concentration – for better mechanistic interpretations and improved risk assessments. Relating effects to internal concentrations should correct for pharmacokinetic differences between species. To show the impact of a proper accounting of cellular concentrations, we calculated

two sets of ECs in the zebrafish embryo: For the first, we use nominal medium concentrations as a measure of dose. For the second, we used the model-predicted embryo concentration at 120 hpf (see Figure 4) as the dose. Because the model is linear with respect to dose, it is possible to obtain and apply a model-computed pharmacokinetic correction factor (f_k) to correct nominal dose ECs. This factor is specific to each compound and can be computed as:

$$f_k = \frac{\text{Nominal concentration}}{\text{Model-predicted embryo concentration}} \quad (35)$$

The corrected ECs are simply obtained as:

$$\text{Corrected EC} = \frac{\text{Nominal dose EC}}{\text{Pharmacokinetic factor}} \quad (36)$$

Table 2 shows the concentrations causing 10, 20 and 50 percent effects (embryo death or malformations, cumulated), uncorrected and corrected for pharmacokinetic at 120 hpf. The data-calibrated model (Figure 3, Table1) was used. Some compounds have high f_k values (almost a factor 5 for 2-methylhexanoic acid and hexanoic acid), but for 4-pentenoic acid, we estimated that no pharmacokinetic correction was necessary. Figure 7 compares the concentrations causing 10% of effects, before and after correction by f_k . We can observe a different ranking between the chemicals. The difference between the minimum and the maximum EC₁₀ value is also wider after pharmacokinetic correction.

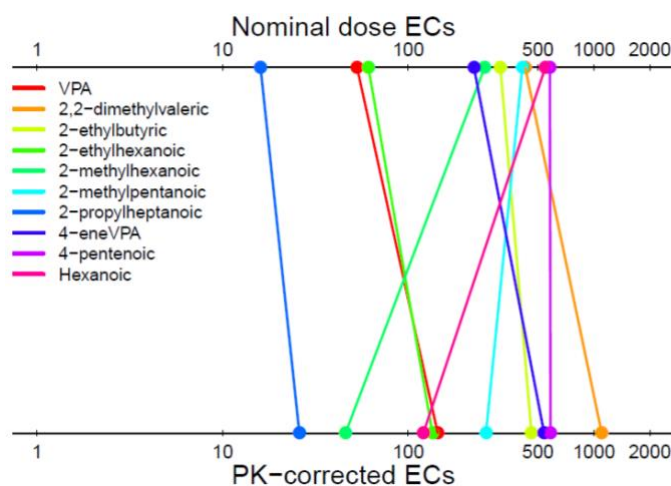


Figure 7. Illustration of the differences between VPA and analogs concentrations (in μM) inducing 10% effects (mortality or malformations, cumulated) when calculated from nominal dose or after correction for pharmacokinetics.

Despite its limitations, the model should be useful for risk assessment. It is simple to use and runs very quickly on a personal computer. For specific malformations or organ toxicity, the *ab initio* predicted tissue/organ concentrations should be used instead of the total embryo concentration. Obviously, if concentration measurements are available to improve the model, they should be used, even if that entails some model fitting. The f_k values we obtained are also a useful summary of the bioaccumulation of VPA and its analogs in the zebrafish embryo. Our results indicate that all the VPA analogs tested accumulate in the embryo and that metabolic clearance is insignificant. For example, for the same medium concentrations, VPA and 2,2-Dimethylvaleric acid lead to the highest concentrations in the embryo (2.5 times the medium concentration), and 2-Methyl hexanoic acid the lowest (21% of the medium concentration) (Table 2). The pharmacokinetic correction should make *EC* more predictive in inter-species extrapolation (however, we did not find suitable animal or human data to confirm this). More importantly, it changes the ranking of the analogs and the order of priorities in a risk assessment context. For example, VPA becoming the 5th most potent (coming from the 2nd rank) and 4-eneVPA going to the 8th rank (from the 4th).

Table 2. Estimated pharmacokinetic correction factor (f_k) and of the 10, 20 and 50% effect concentrations for VPA and its analogs (at 120 hpf) in the zebrafish embryo, uncorrected and corrected for pharmacokinetics (after simultaneous Bayesian calibration Cl_{met} and f_{pc}).

Compound	f_k		EC ₁₀ (μM)		EC ₂₀ (μM)		EC ₅₀ (μM)	
	MPV	IC 95%	Based on	PK	Based on	PK	Based on	PK
			nominal dose	corrected	nominal dose	corrected	nominal dose	corrected
Valproic acid	0.40	[0.34;0.47]	53	133	65	163	96	240
2,2-Dimethylvaleric acid	0.40	[0.36;0.46]	427	1068	445	1113	483	1208
2-Ethylbutyric acid	0.83	[0.62;1.1]	314	378	369	457	510	614
2-Ethylhexanoic acid	1.3	[1.2;1.3]	61	47	71	55	103	79
2-Methylhexanoic acid	4.8	[2.8;7.7]	258	54	280	58	333	69
2-Methylpentanoic acid	1.7	[1.2;2.4]	412	242	425	250	482	284
2-Propylheptanoic acid	0.60	[0.48;0.73]	16	27	17	28	20	33
4-eneVPA	0.54	[0.42;0.74]	226	419	235	435	253	469
4-Pentenoic acid	1.0	[0.72;1.5]	579	579	600	600	645	645
Hexanoic acid	4.5	[4.4;4.8]	548	122	557	124	575	128

4. Conclusions

We developed the structure and equations of the first PBPK model for the zebrafish embryo. Our model integrates previously developed predictive models to account for the impact of physicochemical properties of the chemicals of interest on partition between test system

components, cells of different types, and sub-cellular organelles. It also describes metabolism and anatomical volume changes of the embryo during growth. Its structure and parameter values integrate a large amount of biological information gathered from the scientific literature (on organ volumes *etc.*). However, adult fish physicochemical tissue properties were used, except for the yolk, and future measurements of those parameters in embryos should improve our approximations.

The model can receive several applications. At a first level, using only physicochemical data and QSAR predictions of metabolism, it can be used as a high-throughput *ab initio* tool to make approximate predictions of chemical distribution in different embryonic tissues, as a function of time and exposure levels. At a second-level, when experimental data are available on medium and whole fish or tissue concentrations, the model parameters can be calibrated to estimate *in vivo* metabolism, and make finer predictions. Our model can also be used to guide further research and help design future experiments (for example on the impact of the chorion, effects of transporters on diffusion, *etc.*)

For the chemicals studied here, we found that metabolism was quite low and that the observed kinetics could be reasonably explained simply by volume changes. If this were true on average for most chemicals (which should be checked), the assumption of negligible metabolism could be made, alleviating the need for specific data. That being said, we wish that validated QSAR models were available for the zebrafish embryo.

Yet, our model has been checked with only two sets of concentration time-course data, even if at different exposure levels for the same chemical. Even though it performs rather well at predicting those data, it is certainly not “validated”, and we should not fully trust its predictions: The predictions are afflicted by large uncertainties, which can be estimated by Monte Carlo simulations. More data should be collected to better check and improve the model predictions. We showed how data fitting can improve the model predictions and construe that as an incentive to develop more data. One of the virtues of PBPK models is actually to beg for more data and to direct research questions. The fact that animal PBPK models often have only plasma concentration data to “validate” them does not prevent their extensive use in the pharmaceutical industry, including in regulatory contexts. In any case, our model can be used to relate zebrafish embryo effects to cellular exposures, as demonstrated for VPA analogs. Its use should improve the extrapolation of zebrafish embryo data to human for safety assessment.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 681002 (*EU-ToxRisk*).

References

- [1] A.J. Hill, H. Teraoka, W. Heideman, R.E. Peterson, Zebrafish as a Model Vertebrate for Investigating Chemical Toxicity, *Toxicological Sciences*. 86 (2005) 6–19. <https://doi.org/10.1093/toxsci/kfi110>.
- [2] W. Larisch, K.-U. Goss, Modelling oral up-take of hydrophobic and super-hydrophobic chemicals in fish, *Environmental Science: Processes & Impacts*. (2018). <https://doi.org/10.1039/C7EM00495H>.
- [3] W. Larisch, T.N. Brown, K.-U. Goss, A toxicokinetic model for fish including multiphase sorption features: A high-detailed, physiologically based toxicokinetic model, *Environmental Toxicology and Chemistry*. 36 (2017) 1538–1546. <https://doi.org/10.1002/etc.3677>.
- [4] A. Bernut, G. Lutfalla, L. Kremer, Regard à travers le danio pour mieux comprendre les interactions hôte/pathogène, *Médecine/Sciences*. 31 (2015) 638–646. <https://doi.org/10.1051/medsci/20153106017>.
- [5] J.R. Goldsmith, C. Jobin, Think Small: Zebrafish as a Model System of Human Pathology, *Journal of Biomedicine and Biotechnology*. 2012 (2012) 1–12. <https://doi.org/10.1155/2012/817341>.
- [6] R. Nagel, DarT: The embryo test with the Zebrafish *Danio rerio*--a general model in ecotoxicology and toxicology, *ALTEX*. 19 Suppl 1 (2002) 38–48.
- [7] K. Howe, M.D. Clark, C.F. Torroja, J. Tarrant, C. Berthelot, M. Muffato, J.E. Collins, S. Humphray, K. McLaren, L. Matthews, S. McLaren, I. Sealy, M. Caccamo, C. Churcher, C. Scott, J.C. Barrett, R. Koch, G.-J. Rauch, S. White, W. Chow, B. Kilian, L.T. Quintais, J.A. Guerra-Assunção, Y. Zhou, Y. Gu, J. Yen, J.-H. Vogel, T. Eyre, S. Redmond, R. Banerjee, J. Chi, B. Fu, E. Langley, S.F. Maguire, G.K. Laird, D. Lloyd, E. Kenyon, S. Donaldson, H. Sehra, J. Almeida-King, J. Loveland, S. Trevanion, M. Jones, M. Quail, D. Willey, A. Hunt, J. Burton, S. Sims, K. McLay, B. Plumb, J. Davis, C. Clee, K. Oliver, R. Clark, C. Riddle, D. Elliott, G. Threadgold, G. Harden, D. Ware, B. Mortimer, G. Kerry, P. Heath, B. Phillimore, A. Tracey, N. Corby, M. Dunn, C. Johnson, J. Wood, S. Clark, S. Pelan, G. Griffiths, M. Smith, R. Glithero, P. Howden, N. Barker, C. Stevens, J. Harley, K. Holt, G. Panagiotidis, J. Lovell, H. Beasley, C. Henderson, D. Gordon, K. Auger, D. Wright, J. Collins, C. Raisen, L. Dyer, K. Leung, L. Robertson, K. Ambridge, D. Leongamornlert, S. McGuire, R. Gilderthorp, C. Griffiths, D. Manthavadi, S. Nichol, G. Barker, S. Whitehead, M. Kay, J. Brown, C. Murnane, E. Gray, M. Humphries, N. Sycamore, D. Barker, D. Saunders, J. Wallis, A. Babbage, S. Hammond, M. Mashreghi-Mohammadi, L. Barr, S. Martin, P. Wray, A. Ellington, N. Matthews, M. Ellwood, R. Woodmansey, G. Clark, J. Cooper, A. Tromans, D. Grafham, C. Skuce, R. Pandian, R. Andrews, E. Harrison, A. Kimberley, J. Garnett, N. Fosker, R. Hall, P. Garner, D. Kelly, C. Bird, S. Palmer, I. Gehring, A. Berger, C.M. Dooley, Z. Ersan-Ürün, C. Eser, H. Geiger, M. Geisler, L. Karotki, A. Kirn, J. Konantz, M. Konantz, M. Oberländer, S. Rudolph-Geiger, M. Teucke, K. Osoegawa, B. Zhu, A. Rapp, S. Widaa, C. Langford, F. Yang, N.P. Carter, J. Harrow, Z. Ning, J. Herrero, S.M.J. Searle, A. Enright, R. Geisler, R.H.A. Plasterk, C. Lee, M. Westerfield, P.J. de Jong, L.I. Zon, J.H. Postlethwait, C. Nüsslein-Volhard, T.J.P. Hubbard,

- H.R. Crollius, J. Rogers, D.L. Stemple, The zebrafish reference genome sequence and its relationship to the human genome, *Nature*. 496 (2013) 498–503. <https://doi.org/10.1038/nature12111>.
- [8] C.S. Martinez, D.A. Feas, M. Siri, D.E. Igartúa, N.S. Chiaramoni, S. del V. Alonso, M.J. Prieto, In vivo study of teratogenic and anticonvulsant effects of antiepileptics drugs in zebrafish embryo and larvae, *Neurotoxicology and Teratology*. 66 (2018) 17–24. <https://doi.org/10.1016/j.ntt.2018.01.008>.
- [9] M. Ekker, M.-A. Akimento, Le poisson zèbre (*danio rerio*) un modèle en biologie du développement, *Médecine/Sciences*. 7 (1991) 553–560.
- [10] B. Kais, R. Ottermanns, F. Scheller, T. Braunbeck, Modification and quantification of in vivo EROD live-imaging with zebrafish (*Danio rerio*) embryos to detect both induction and inhibition of CYP1A, *Science of The Total Environment*. 615 (2018) 330–347. <https://doi.org/10.1016/j.scitotenv.2017.09.257>.
- [11] U. Strähle, S. Scholz, R. Geisler, P. Greiner, H. Hollert, S. Rastegar, A. Schumacher, I. Selderslaghs, C. Weiss, H. Witters, T. Braunbeck, Zebrafish embryos as an alternative to animal experiments—A commentary on the definition of the onset of protected life stages in animal welfare regulations, *Reproductive Toxicology*. 33 (2012) 128–132. <https://doi.org/10.1016/j.reprotox.2011.06.121>.
- [12] N. Quignot, J. Hamon, F.Y. Bois, Extrapolating *in vitro* results to predict human toxicity, in: A. Bal-Price, P. Jennings (Eds.), *In Vitro Toxicology Systems*, Springer Science, New-York, 2014: pp. 531–550.
- [13] P.F. Landrum, M.J. Lydy, H. Lee, Toxicokinetics in aquatic systems: Model comparisons and use in hazard assessment, *Environmental Toxicology and Chemistry*. 11 (1992) 1709–1725. <https://doi.org/10.1002/etc.5620111205>.
- [14] E.S. Salmina, D. Wondrousch, R. Kühne, V.A. Potemkin, G. Schüürmann, Variation in predicted internal concentrations in relation to PBPK model complexity for rainbow trout, *Science of The Total Environment*. 550 (2016) 586–597. <https://doi.org/10.1016/j.scitotenv.2016.01.107>.
- [15] K. Krishnan, T. Peyret, Physiologically Based Toxicokinetic (PBTK) Modeling in Ecotoxicology, in: J. Devillers (Ed.), *Ecotoxicology Modeling*, Springer US, Boston, MA, 2009: pp. 145–175. https://doi.org/10.1007/978-1-4419-0197-2_6.
- [16] J. Feng, Y. Gao, M. Chen, X. Xu, M. Huang, T. Yang, N. Chen, L. Zhu, Predicting cadmium and lead toxicities in zebrafish (*Danio rerio*) larvae by using a toxicokinetic–toxicodynamic model that considers the effects of cations, *Science of The Total Environment*. 625 (2018) 1584–1595. <https://doi.org/10.1016/j.scitotenv.2018.01.068>.
- [17] M. Khazaei, C.A. Ng, Evaluating parameter availability for physiologically based pharmacokinetic (PBPK) modeling of perfluorooctanoic acid (PFOA) in zebrafish, *Environmental Science: Processes & Impacts*. 20 (2018) 105–119. <https://doi.org/10.1039/C7EM00474E>.
- [18] A.R.R. Péry, J. Devillers, C. Brochot, E. Mombelli, O. Palluel, B. Piccini, F. Brion, R. Beaudouin, A Physiologically Based Toxicokinetic Model for the Zebrafish *Danio rerio*, *Environmental Science & Technology*. 48 (2014) 781–790. <https://doi.org/10.1021/es404301q>.

- [19] A. Grech, C. Tebby, C. Brochot, F.Y. Bois, A. Bado-Nilles, J.-L. Dorne, N. Quignot, R. Beaudouin, Generic physiologically-based toxicokinetic modelling for fish: Integration of environmental factors and species variability, *Science of The Total Environment*. 651 (2019) 516–531. <https://doi.org/10.1016/j.scitotenv.2018.09.163>.
- [20] M. Brinkmann, C. Schlechtriem, M. Reininghaus, K. Eichbaum, S. Buchinger, G. Reifferscheid, H. Hollert, T.G. Preuss, Cross-species extrapolation of uptake and disposition of neutral organic chemicals in fish using a multispecies physiologically-based toxicokinetic model framework, *Environmental Science and Technology*. 50 (2016) 1914–1923. <https://doi.org/10.1021/acs.est.5b06158>.
- [21] S. Brox, B. Seiwert, E. Küster, T. Reemtsma, Toxicokinetics of Polar Chemicals in Zebrafish Embryo (*Danio rerio*): Influence of Physicochemical Properties and of Biological Processes, *Environmental Science & Technology*. 50 (2016) 10264–10272. <https://doi.org/10.1021/acs.est.6b04325>.
- [22] F.A.P.C. Gobas, X. Zhang, Measuring bioconcentration factors and rate constants of chemicals in aquatic organisms under conditions of variable water concentrations and short exposure time, *Chemosphere*. 25 (1992) 1961–1971. [https://doi.org/10.1016/0045-6535\(92\)90035-P](https://doi.org/10.1016/0045-6535(92)90035-P).
- [23] J.C. Otte, B. Schultz, D. Fruth, E. Fabian, B. van Ravenzwaay, B. Hidding, E.R. Salinas, Intrinsic xenobiotic metabolizing enzyme activities in early life stages of zebrafish (*Danio rerio*), *Toxicological Sciences*. 159 (2017) 86–93. <https://doi.org/10.1093/toxsci/kfx116>.
- [24] C. Fisher, S. Siméon, M. Jamei, I. Gardner, Y.F. Bois, VIVD: Virtual in vitro distribution model for the mechanistic prediction of intracellular concentrations of chemicals in in vitro toxicity assays, *Toxicology in Vitro*. 58 (2019) 42–50. <https://doi.org/10.1016/j.tiv.2018.12.017>.
- [25] K. Fathe, A. Palacios, R.H. Finnell, Brief report novel mechanism for valproate-induced teratogenicity: Novel Mechanism for Valproate-Induced Teratogenicity, *Birth Defects Research Part A: Clinical and Molecular Teratology*. 100 (2014) 592–597. <https://doi.org/10.1002/bdra.23277>.
- [26] C.-M. Chuang, C.-H. Chang, H.-E. Wang, K.-C. Chen, C.-C. Peng, C.-L. Hsieh, R.Y. Peng, Valproic Acid Downregulates RBP4 and Elicits Hypervitaminosis A-Teratogenesis—A Kinetic Analysis on Retinol/Retinoic Acid Homeostatic System, *PLoS ONE*. 7 (2012) e43692. <https://doi.org/10.1371/journal.pone.0043692>.
- [27] C.J. Phiel, F. Zhang, E.Y. Huang, M.G. Guenther, M.A. Lazar, P.S. Klein, Histone Deacetylase Is a Direct Target of Valproic Acid, a Potent Anticonvulsant, Mood Stabilizer, and Teratogen, *Journal of Biological Chemistry*. 276 (2001) 36734–36741. <https://doi.org/10.1074/jbc.M101287200>.
- [28] T. Braunbeck, E. Lammer, Fish Embryo Toxicity Assays (UBA contract number 20385422), (2006).
- [29] E. Lammer, G.J. Carr, K. Wendler, J.M. Rawlings, S.E. Belanger, Th. Braunbeck, Is the fish embryo toxicity test (FET) with the zebrafish (*Danio rerio*) a potential alternative for the fish acute toxicity test?, *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*. 149 (2009) 196–209. <https://doi.org/10.1016/j.cbpc.2008.11.006>.

- [30] T. Braunbeck, B. Kais, E. Lammer, J. Otte, K. Schneider, D. Stengel, R. Strecker, The fish embryo test (FET): origin, applications, and future, *Environmental Science and Pollution Research*. 22 (2015) 16247–16261. <https://doi.org/10.1007/s11356-014-3814-7>.
- [31] M.R. Embry, S.E. Belanger, T.A. Braunbeck, M. Galay-Burgos, M. Halder, D.E. Hinton, M.A. Léonard, A. Lillicrap, T. Norberg-King, G. Whale, The fish embryo toxicity test as an animal alternative method in hazard and risk assessment and scientific research, *Aquatic Toxicology*. 97 (2010) 79–87. <https://doi.org/10.1016/j.aquatox.2009.12.008>.
- [32] ISO, International Organization for Standardization. Water quality - Determination of the 28 acute lethal toxicity of substances to a freshwater fish [*Brachydanio rerio* Hamilton-Buchanan 29 (Teleostei, Cyprinidae)]. ISO 7346-3: Flow-through method. Available: [<http://www.iso.org>]., (1996).
- [33] OECD, OECD guidelines for the testing of chemicals. Fish Embryo Acute Toxicity (FET) Test, (2013). https://www.oecd-ilibrary.org/environment/test-no-236-fish-embryo-acute-toxicity-fet-test_9789264203709-en.
- [34] T. Braunbeck, M. Boettcher, H. Hollert, T. Kosmehl, E. Lammer, E. Leist, M. Rudolf, N. Seitz, Towards an alternative for the acute fish LC(50) test in chemical assessment: the fish embryo toxicity test goes multi-species -- an update, *ALTEX*. 22 (2005) 87–102.
- [35] H. Hollert, S. Keiter, N. König, M. Rudolf, M. Ulrich, T. Braunbeck, A new sediment contact assay to assess particle-bound pollutants using zebrafish (*danio rerio*) embryos, *Journal of Soils and Sediments*. 3 (2003) 197–207. <https://doi.org/10.1065/jss2003.09.085>.
- [36] C.B. Kimmel, W.W. Ballard, S.R. Kimmel, B. Ullmann, T.F. Schilling, Stages of embryonic development of the zebrafish, *Developmental Dynamics*. 203 (1995) 253–310. <https://doi.org/10.1002/aja.1002030302>.
- [37] N. Hachicho, S. Reithel, A. Miltner, H.J. Heipieper, E. Küster, T. Luckenbach, Body mass parameters, lipid profiles and protein contents of zebrafish embryos and effects of 2,4-dinitrophenol exposure, *PLoS One*. 10 (2015) e0134755. <https://doi.org/10.1371/journal.pone.0134755>.
- [38] M. Hagedorn, F.W. Kleinhans, R. Freitas, J. Liu, E.W. Hsu, D.E. Wildt, W.F. Rall, Water distribution and permeability of zebrafish embryos, *Brachydanio rerio*, *The Journal of Experimental Zoology*. 278 (1997) 356–371. [https://doi.org/10.1002/\(sici\)1097-010x\(19970815\)278:6<356::aid-jez3>3.0.co;2-n](https://doi.org/10.1002/(sici)1097-010x(19970815)278:6<356::aid-jez3>3.0.co;2-n).
- [39] P. Bernillon, F.Y. Bois, Statistical issues in toxicokinetic modeling: a Bayesian perspective, *Environmental Health Perspectives*. 108 (suppl. 5) (2000) 883–893.
- [40] F.Y. Bois, Bayesian inference, in: B. Reisfeld, A.N. Mayeno (Eds.), *Computational Toxicology Vol. II*, Humana Press, New-York, 2012: pp. 597–636.
- [41] A.F.M. Smith, G.O. Roberts, Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods, *Journal of the Royal Statistical Society Series B*. 55 (1993) 3–23.
- [42] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences (with discussion), *Statistical Science*. 7 (1992) 457–511.
- [43] J.W. Nichols, J.M. McKim, M.E. Andersen, M.L. Gargas, H.J. Clewell, R.J. Erickson, A physiologically based toxicokinetic model for the uptake and disposition of waterborne organic chemicals in fish, *Toxicol. Appl. Pharmacol.* 106 (1990) 433–447.

- [44] R Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria., 2013. <http://www.R-project.org>.
- [45] F.Y. Bois, GNU MCSim: Bayesian statistical inference for SBML-coded systems biology models, *Bioinformatics*. 25 (2009) 1453–1454. <https://doi.org/10.1093/bioinformatics/btp162>.
- [46] J.M.Z. Comenges, E. Joossens, J.V.S. Benito, A. Worth, A. Paini, Theoretical and mathematical foundation of the Virtual Cell Based Assay – A review, *Toxicology in Vitro*. 45 (2017) 209–221. <https://doi.org/10.1016/j.tiv.2016.07.013>.
- [47] E. Papa, L. van der Wal, J.A. Arnot, P. Gramatica, Metabolic biotransformation half-lives in fish: QSAR modeling and consensus analysis, *Science of The Total Environment*. 470–471 (2014) 1040–1046. <https://doi.org/10.1016/j.scitotenv.2013.10.068>.
- [48] J.A. Arnot, W. Meylan, J. Tunkel, P.H. Howard, D. Mackay, M. Bonnell, R.S. Boethling, A quantitative structure–activity relationship for predicting metabolic biotransformation rates for organic chemicals in fish, *Environ Toxicol Chem*. 28 (2009) 1168. <https://doi.org/10.1897/08-289.1>.
- [49] S. Fischer, N. Klüver, K. Burkhardt-Medicke, M. Pietsch, A.-M. Schmidt, P. Wellner, K. Schirmer, T. Luckenbach, Abcb4 acts as multixenobiotic transporter and active barrier against chemical uptake in zebrafish (*Danio rerio*) embryos, *BMC Biology*. 11 (2013) 69. <https://doi.org/10.1186/1741-7007-11-69>.

Corrigendum to “Development of a generic zebrafish embryo PBPK model and application to the developmental toxicity assessment of valproic acid analogs” [Reprod. Toxicol. 2020, 93:219-229]

Ségolène Siméon[†], Katharina Brotzmann[‡], Ciaran Fisher[§], Iain Gardner[§], Steve Silvester[‡], Richard MacLennan[‡], Paul Walker[‡], Thomas Braunbeck[‡] and Frederic Y. Bois^{*,§}.

[†] INERIS, METO unit, Parc ALATA BP2, Verneuil en Halatte, France.

[‡] University of Heidelberg, Aquatic Ecology and Toxicology, Centre for Organismal Studies (COS), Im Neuenheimer Feld 504, D-69120 Heidelberg, Germany.

[§] CERTARA UK Limited, Simcyp Division, Level 2-Acero, 1 Concourse Way, Sheffield, S1 2BJ, United Kingdom; Email: *frederic.bois@certara.com

^{||} Cyprotex Discovery Ltd., No. 24 Mereside, Alderley Park, Macclesfield, Cheshire, SK10 4TG, United Kingdom.

During the experiments reported in our paper, according to the OECD protocol for the acute toxicity test on zebrafish embryos, the water pH was checked only at the start of the assay. However, the pH buffer used was not strong enough to keep pH constant when adding the various acidic molecules assayed –valproic acid (VPA) and nine analogs. After independent pH measurements, a water pH decrease was noted with increasing concentrations of the test chemicals (Table 1). However, when analysing the embryo concentration data, we mistakenly considered the medium pH to have a constant buffered value of 7.4. Since pH has a strong effect on the ionisation of test compounds in aqueous solution, it can have also a strong impact of the accumulation and distribution of those compounds in the body of the embryos. We have therefore redone all affected analyses (simulations of the data published by Brox *et al.*, except for valproic acid, were not affected), and we give here the updated results. Fish tissues were considered sufficiently buffered at their normal pH in the range of compound concentrations used.

Table 1. Water pH value according to the nominal concentration for each compound assayed.

Compound	Nominal concentration (μM)	Measured pH
Valproic acid (VPA)	6.25	7.64
	12.5	7.55
	25	7.41
	50	7.32
	100	7.11
	200	6.83
	400	6.23
2,2-Dimethylvaleric acid	312.5	7.05
	500	6.71
	800	5.88
2-Ethylbutyric acid	100	7.43
	200	7.11
	400	6.64
	800	5.36
2-Ethylhexanoic acid	12.5	7.52
	25	7.35
	50	7.18
	100	7.08
	200	6.72
2-Methylhexanoic acid	125	7.42
	250	7.17
	500	7.03
2-Methylpentanoic acid	177.7	7.10
	266.6	6.97
	400	6.56
	600	6.06
2-Propylheptanoic acid	12.5	7.67
	25	7.63
	50	7.48
4-ene VPA	200	7.70
	400	7.42
4-Pentenoic acid	500	6.78
	1000	5.44
Hexanoic acid	520	6.43
	625	6.27

To determine the partition coefficient of each compound at each concentration, we used the VIVD sub-model, setting the water medium pH at the correct values. Figure 1 shows the impact of pH on the VPA organs over medium unbound partition coefficients. The air and plastic to medium unbound partition coefficients are not influenced by pH.

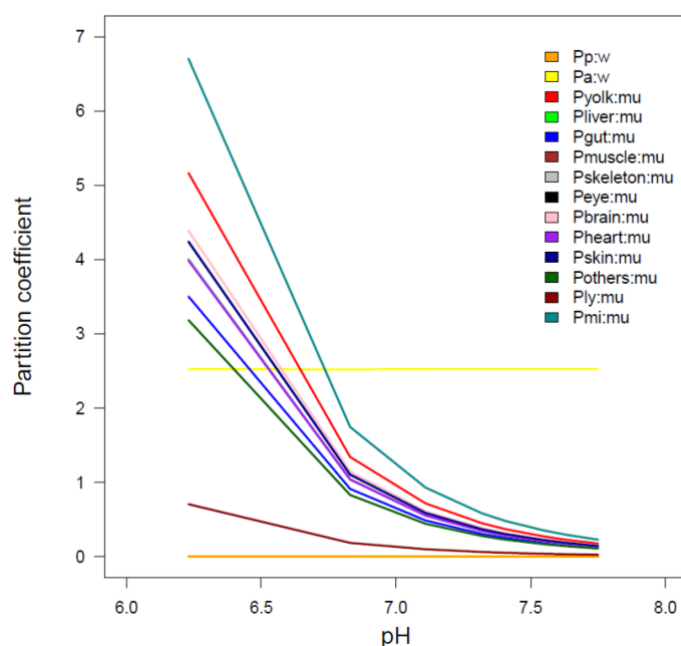


Figure 1. Partition coefficients of valproic acid as a function of water pH.

With the corrected external pH and partition coefficient, the new *ab initio* predictions of the experimental data on all VPA analogs (before parameter fitting) are shown in Figure 2. The correction for external pH markedly improves the predictions with almost all compounds in the ten-fold error interval.

For each VPA analog, we re-estimated metabolic clearance (Cl_{met}) and the adjustment factor of the partition coefficients (f_{pc}), taking actual pH values into account. Figure 3 shows observed vs. correctly predicted concentrations in total embryo and culture medium. Overall, the fit is improved. Hexanoic acid, VPA, 4-eneVPA and 2-methylhexanoic acid are now the four worst predicted chemicals and some of their data points do not fall in the three-fold error interval. The medium concentrations are still slightly under-predicted by the model with a median relative error of 0.76. The total embryo concentrations are well predicted with a median relative error of 1.06. The best estimate of the residual error (σ) corresponds on average to a factor 2.07, lower than the one previously one obtained when using the wrong pH.

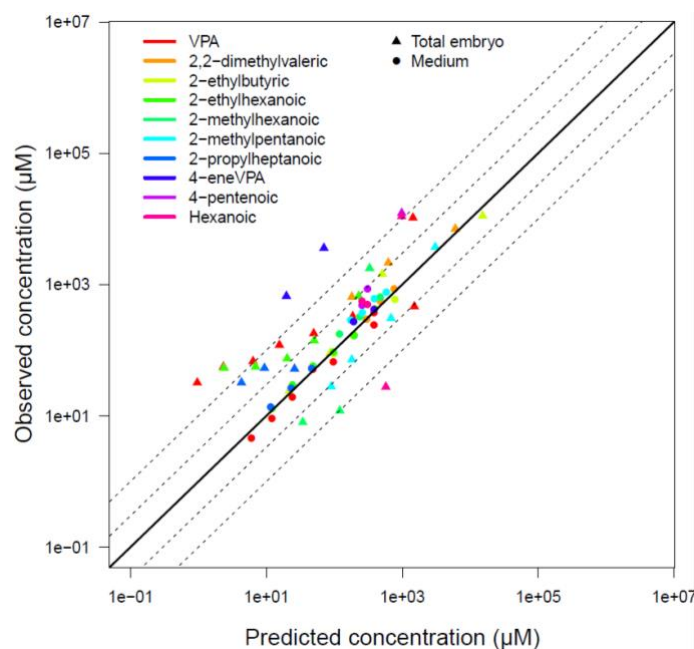


Figure 2. Observed versus model-predicted concentrations in the zebrafish embryo and in culture medium for valproic acid and nine analogs. The VIVD-computed parameter values were used without further adjustment. The black line corresponds to perfect predictions.

Dashed lines delineate the three- and ten-fold error bounds.

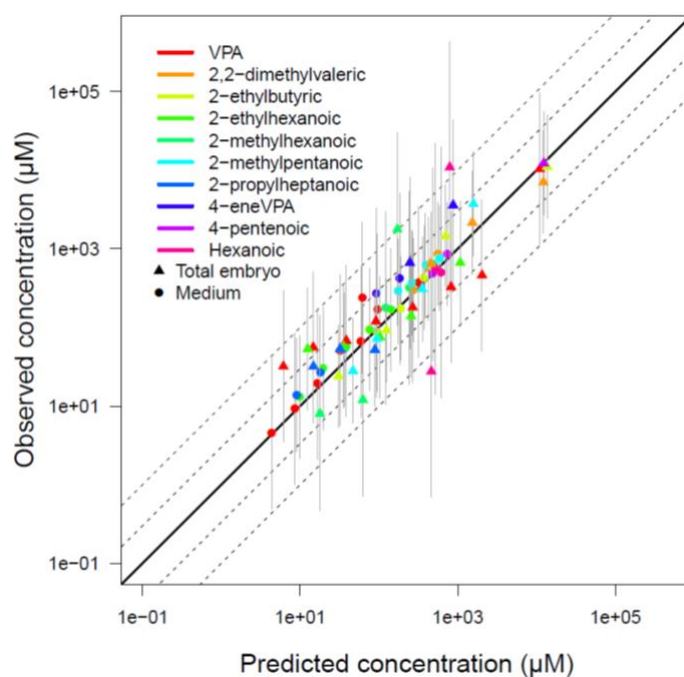


Figure 3. Observed *versus* predicted concentrations of valproic acid (VPA) and nine analogs, in the zebrafish embryo and culture medium. This is the best fit obtained after Cl_{met} and f_{pc} were simultaneously estimated. The solid black line corresponds to perfect fit. Dashed lines correspond to the three- and ten-fold error intervals. The grey bars correspond to $\pm$ one residual SD (σ).

Figure 4 presents the correct observed and predicted concentrations of VPA in medium and in total embryo as a function of time, following exposure to various concentrations, after simultaneous estimation of Cl_{met} and f_{pc} . The concentrations were normalized by the corresponding nominal concentration. The Figure shows, for each exposure concentration, 10 posterior predictions obtained using random parameters to illustrate uncertainty. The model under-predicts medium concentrations at highest exposures but gives reasonably good predictions of total embryo concentrations.

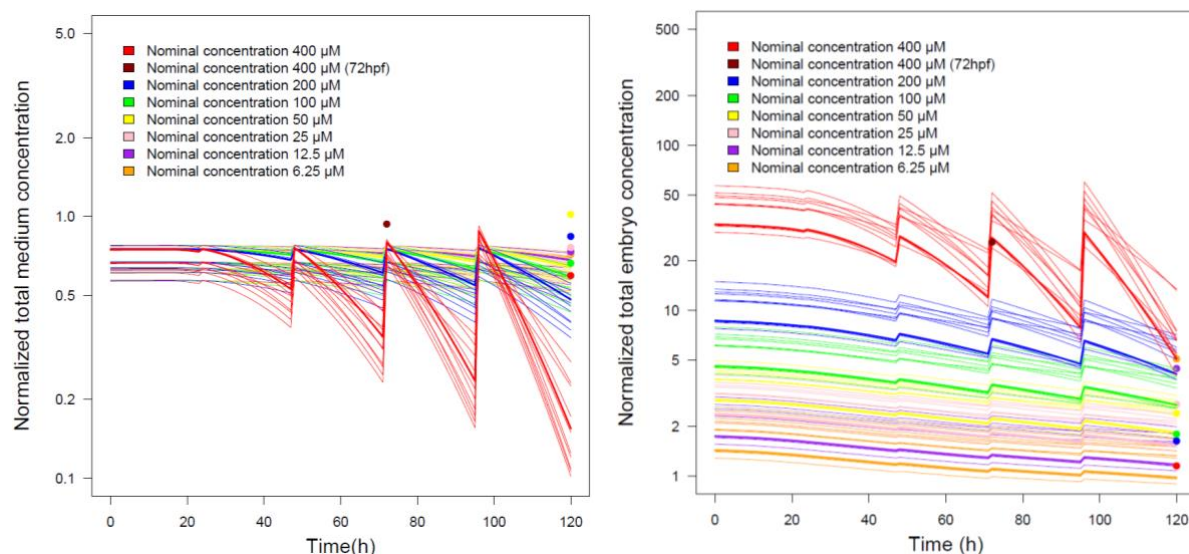


Figure 4. Predicted (lines) and observed (dots) VPA concentrations in medium (left) and in total embryo (right) as a function of time, after estimating Cl_{met} and f_{pc} . All concentrations were normalized to their respective nominal concentration for plotting. Culture medium was replaced every day. The thick lines are the maximum posterior predicted concentration time-course. The thin lines are 10 predictions obtained using random parameter vectors drawn from their posterior distribution.

Table 2 summarizes the corrected prior distributions used for Cl_{met} and f_p . Table 3 summarizes, for each chemical, the posterior distributions of Cl_{met} , f_{pc} and uncertainty SD σ obtained by MCMC calibration of the model with the concentration data. For the various compounds, the metabolic clearance best estimates were between 10^{-12} to 10^{-13} L/min. So, the conclusion of negligible metabolism is still valid. The best estimates of the partition coefficient scaling factor, f_{pc} , are in the range of 0.4 (for Hexanoic acid) to 26 for 4-eneVPA. According to their 95% confidence intervals, the f_{pc} values for 2-Ethylbutyric acid, 2-Methylhexanoic acid, 2-Methylpentanoic acid, 4-Pentenoic acid, and Hexanoic acid are not significantly different from 1 (implying that a correction factor is not really needed for those compounds).

Table 2. Prior distribution of CL_{met} and f_{pc} for each compound.

Compound	Uniform prior on CL_{met}	Uniform prior on f_{pc}
Valproic acid	$0 - 1 \times 10^{-11}$	$0 - 20$
2,2- Dimethylvaleric acid	$0 - 2 \times 10^{-11}$	$0 - 5$
2-Ethylbutyric acid	$0 - 1 \times 10^{-11}$	$0 - 5$
2-Ethylhexanoic acid	$0 - 1 \times 10^{-11}$	$0 - 20$
2-Methylhexanoic acid	$0 - 5 \times 10^{-10}$	$0 - 5$
2-Methylpentanoic acid	$0 - 1 \times 10^{-11}$	$0 - 10$
2-Propylheptanoic acid	$0 - 1 \times 10^{-11}$	$0 - 10$
4-eneVPA	$0 - 1 \times 10^{-10}$	$0 - 200$
4-Pentenoic acid	$0 - 1 \times 10^{-11}$	$0 - 20$
Hexanoic acid	$0 - 7 \times 10^{-11}$	$0 - 10$

Table 3. Estimation of means, standard deviation (SD), 95% confidence intervals (IC95%) and maximum posterior (MP) value for the metabolic clearance, the partition coefficient correction factor, and residual uncertainty SD σ , for VPA and its analogs.

Compound	CL_{met} (L/min)			f_{pc}			σ		
	Mean $\pm$ SD	IC 95%	MP	Mean $\pm$ SD	IC 95%	MP	Mean $\pm$ SD	IC 95%	MP
Valproic acid	$4.3 \times 10^{-12} \pm 1.8 \times 10^{-12}$	$[1.1 \times 10^{-12}; 8.0 \times 10^{-12}]$	4.2×10^{-12}	9.7 ± 3.5	$[4.3; 18]$	8.8	2.6 ± 0.45	$[1.9; 3.7]$	2.2
2,2- Dimethylvaleric acid	$1.5 \times 10^{-12} \pm 1.3 \times 10^{-12}$	$[5.8 \times 10^{-14}; 5.0 \times 10^{-12}]$	7.1×10^{-13}	3.0 ± 1.0	$[1.3; 4.8]$	2.7	2 ± 0.48	$[1.5; 3.2]$	1.5
2-Ethylbutyric acid	$1.2 \times 10^{-12} \pm 8.2 \times 10^{-13}$	$[1.1 \times 10^{-13}; 3.1 \times 10^{-12}]$	9.9×10^{-13}	1.6 ± 0.57	$[0.79; 3.1]$	1.5	1.8 ± 0.36	$[1.5; 2.9]$	1.5
2-Ethylhexanoic acid	$3.5 \times 10^{-12} \pm 2.4 \times 10^{-12}$	$[1.9 \times 10^{-13}; 8.9 \times 10^{-12}]$	1.2×10^{-12}	8.4 ± 3.4	$[3.5; 17]$	6.4	2.2 ± 0.44	$[1.6; 3.3]$	1.8
2-Methylhexanoic acid	$6.8 \times 10^{-11} \pm 7.4 \times 10^{-11}$	$[1.5 \times 10^{-12}; 2.8 \times 10^{-10}]$	4.1×10^{-13}	0.93 ± 0.78	$[0.16; 3.3]$	0.51	3.6 ± 0.82	$[2.4; 5.5]$	2.8
2-Methylpentanoic acid	$3.6 \times 10^{-12} \pm 2.6 \times 10^{-12}$	$[1.2 \times 10^{-13}; 9.2 \times 10^{-12}]$	1.1×10^{-13}	0.62 ± 0.32	$[0.27; 1.5]$	0.51	2.1 ± 0.51	$[1.5; 3.5]$	1.5
2-Propylheptanoic acid	$4.0 \times 10^{-12} \pm 2.9 \times 10^{-12}$	$[1.6 \times 10^{-13}; 9.6 \times 10^{-12}]$	1.8×10^{-13}	5.7 ± 2.0	$[2.1; 9.6]$	4.4	2.2 ± 0.54	$[1.6; 3.6]$	1.7
4-eneVPA	$1 \times 10^{-11} \pm 1.2 \times 10^{-11}$	$[1.5 \times 10^{-13}; 4.5 \times 10^{-11}]$	3.7×10^{-13}	51 ± 40	$[7.5; 165]$	26	3.5 ± 0.9	$[2.2; 5.5]$	2.5
4-Pentenoic acid	$1.9 \times 10^{-12} \pm 1.8 \times 10^{-12}$	$[8.8 \times 10^{-14}; 7.2 \times 10^{-12}]$	1.0×10^{-12}	1.8 ± 1.7	$[0.26; 7.0]$	0.99	2.3 ± 0.74	$[1.5; 4.2]$	1.5
Hexanoic acid	$1.9 \times 10^{-11} \pm 1.6 \times 10^{-11}$	$[7.0 \times 10^{-13}; 5.9 \times 10^{-11}]$	1.2×10^{-13}	1.3 ± 1.6	$[0.10; 7.0]$	0.4	4.2 ± 0.85	$[2.8; 6.1]$	3.7

Figure 5 shows our corrected predictive simulations of Brox *et al.* data on VPA, using the best parameter estimates for Cl_{met} and f_{pc} . Here also the predictions are much improved.

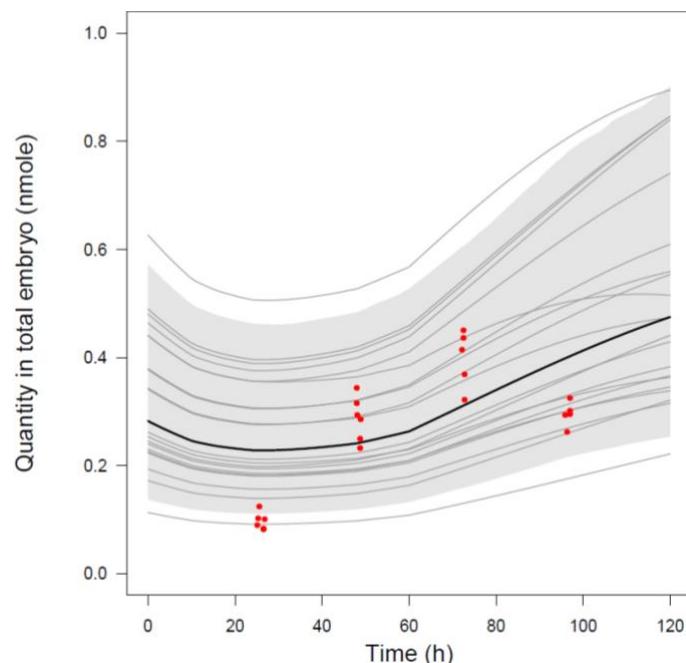


Figure 5. Predicted (lines) and data by Brox *et al.* (dots). VPA quantities in total embryo as a function of time, after estimating Cl_{met} and f_{pc} from our data. The culture medium was not replaced. The grey area defines the 95% confidence interval. The thick black line is the maximum posterior prediction. The thin lines are 20 predictions obtained using random parameter vectors drawn from their posterior distribution.

Finally, Table 4 shows the concentrations causing 10 percent effect (EC_{10}) obtained using nominal concentrations and those obtained using correctly predicted internal concentration at 120 hpf. The model is no longer linear since partition coefficients do not vary linearly with exposure concentration (Figure 1). Therefore, we cannot use a simple pharmacokinetic factor as before. However, we can still compute, for each nominal EC_{10} , an internal EC_{10} by running the model with the nominal EC_{10} as medium concentration, the measured pH, and all other relevant conditions. Figure 6 compares the nominal EC_{10} to the predicted internal EC_{10} (better suited for inter-species transposition). We still observe different rankings between the two kinds of estimates. The PK-corrected EC_{10} values tend to be higher than the nominal EC_{10} values.

Table 4. EC₁₀ values from nominal and internal concentrations (μM)

Compound	Nominal EC ₁₀ (μM)	Total embryo EC ₁₀ (μM)
VPA	53	99
2,2-Dimethylvaleric acid	427	987
2-Ethylbutyric acid	314	362
2-Ethylhexanoic acid	61	132
2-Methylhexanoic acid	258	65
2-Methylpentanoic acid	412	391
2-Propylheptanoic acid	16	19
4-ene VPA	226	302
4-Pentenoic acid	579	868
Hexanoic acid	548	380

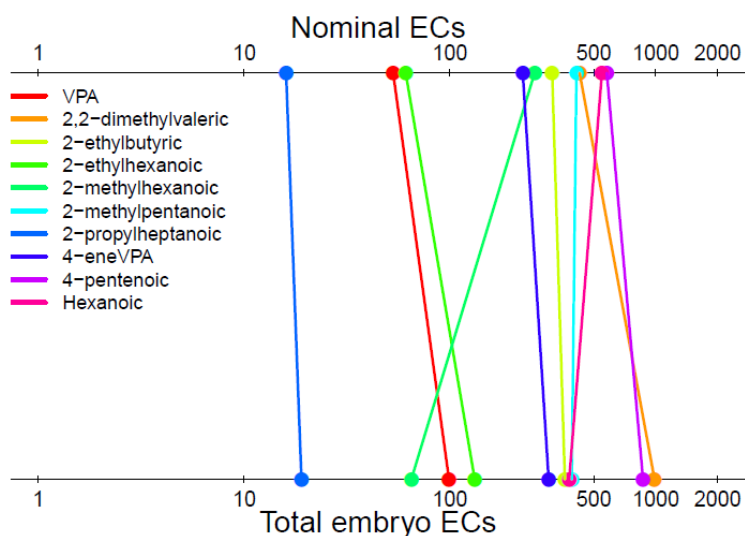


Figure 6. Illustration of the differences between VPA and analogs concentrations (in μM) inducing 10% effects (mortality or malformations, cumulated) when calculated from nominal concentration or from predicted internal concentration.

Over all, using the correct pH values improves all model predictions. Metabolism appears to be still negligible.

Two typographic errors were also spotted: The partition coefficients are “organs over medium” partition coefficients: $P_{mu:i}$ has to be replaced by $P_{i:mu}$, and Equation 1 should be:

$$C_i = C_{medium,u} \times f_{pc} \times P_{i:mu} \quad (1)$$

This also applies to Equation 2, where f_{uu} has to be replaced by f_{ui} :

$$C_{medium,u} = \frac{Q_{parent}}{\frac{V_{medium}}{f_{ui}} + P_{a:w} \times f_{ui} \times V_{air}(t) + P_{p:w} \times S_{medium}(t) + \sum(P_{mu:j} \times V_j(t))} \quad (2)$$

The authors would like to apologize for any inconvenience caused.

CHAPITRE IV. Modèle PK-PD multi-états pour étudier les effets des composés chimiques

1. Un nouveau modèle multi-états de la toxicité développementales des composés chimiques chez l'embryon de poisson zèbre

1.1. Objectif des travaux et contribution

Pour les essais de toxicité développementale chez l'embryon de poisson zèbre, les directives de l'OCDE recommandent d'effectuer une analyse dose-réponse distincte pour chaque heure d'observation. Cependant, les événements à un moment donné sont en partie conditionnés par les phénomènes passés. La séparation des données par groupe de temps réduit également le nombre de points de données par analyse, les rendant obsolètes ou les affectant en réduisant leur puissance d'information. L'analyse standard des données de temps d'occurrence des malformations utilise des modèles de survie, couramment utilisés dans les études cliniques ou dans les analyses de temps défaillant. De plus, les occurrences de malformations ne sont pas indépendantes et constituent un phénomène de risques concurrentiels. La solution généralement recommandée pour résoudre ces problèmes de survie complexes est les modèles multi-états.

Dans le cadre du projet européen EU-ToxRisk, les tests de toxicité suivent les directives stipulées par l'OCDE. Nous avons donc décidé de contribuer à l'amélioration des analyses de données de toxicité chez l'embryon de poisson zèbre en développant un modèle de survie multi-états adaptés au test de l'OCDE. Il a été construit dans le cadre formel des modèles multi-états Markoviens, puis couplé au modèle PBPK décrit précédemment dans le Chapitre 3 pour générer un modèle tenant compte de la variation de concentration de VPA dans l'embryon et capable d'intégrer l'ensemble des données concentration-temps-effets collectées dans un protocole expérimental OCDE standard, au cours des cinq premiers jours du développement d'un embryon. Par ajustement bayésien aux données de malformations, les paramètres des taux de transitions d'un état à l'autre ont été estimés. Ainsi, nous avons pu déterminer les probabilités de transition résultantes, et prédire les concentrations effectives associées.

Ces travaux de recherche ont été présentés dans l'article « *Multistate Models of Developmental Toxicity: Application to Valproic Acid-Induced Malformations in the Zebrafish*

Embryo », soumis dans la revue *Toxicology and Applied Pharmacology* en 2020 et présenté ci-après. Les informations supplémentaires sont disponibles en Annexe 3.

1.2. Articulation avec le modèle PBPK précédent

Lors de l'analyse expérimentale dose-réponse suite à l'exposition à un composé, l'apparition des effets est reliée à la concentration nominale testée. Cependant, les malformations observées sont le résultat du comportement du composé au sein du tissu concerné. Il est donc plus pertinent de les attribuer à la concentration interne. De plus, pour extrapoler l'évaluation de la toxicité du poisson zèbre à l'homme, il est essentiel de considérer les différences pharmacocinétiques entre les deux espèces. Le modèle PBPK présenté précédemment en Chapitre 3 permet d'aborder cette problématique par prédiction de la concentration interne. Ainsi, nous avons développé le modèle multi-états en intégrant le profil pharmacocinétique de la concentration interne prédite du VPA pour associer l'occurrence des malformations au cours du temps au comportement de la molécule dans l'embryon de poisson zèbre.

1.3. Perspectives

Le modèle PK-PD multi-états proposé pour étudier la toxicité développementale est applicable à de nombreux produits chimiques et à de nombreuses espèces pour l'évaluation de risques (éco)-toxicologiques. Il est général et représente un travail préliminaire à l'entière interprétation des diverses malformations observées. Il serait possible de différencier les effets à différentes étapes du processus de malformation, inaccessibles par des analyses dose-réponse simples. Par conception optimale ou calculs statistiques de la taille des échantillons, le modèle multi-états pourrait également permettre de réduire l'utilisation des animaux pour les tests de toxicité. Il pourrait être étendu en considérant plus d'états et en différenciant la sévérité des malformations. Son couplage à un modèle PBPK deviendrait alors plus pertinent puisque les concentrations tissulaires spécifiques seraient reliées à des malformations organes spécifiques. L'approche est relativement complexe et devrait être vérifiée avec un large panel de produits chimiques aux propriétés diverses avant de pouvoir être utilisée en routine. Cependant, le modèle devrait aider les communautés scientifiques et d'évaluation des risques à améliorer l'analyse des données.

2. Résumé et article du modèle PK-PD multi-états

Pour la détermination de la toxicité aiguë des produits chimiques chez les embryons de poisson zèbre (*Danio rerio*), la ligne directrice 236 de l'OCDE prévoit une analyse dose-réponse de quatre effets létaux et une caractérisation quantitative des anomalies, y compris leur dépendance au temps. En routine, les données sont analysées séparément à différents temps d'observation. Cependant, les observations à temps fixe dépendent fortement des effets antérieurs et doivent être analysées conjointement. Pour résoudre ce problème, nous avons développé un modèle multi-états pour déterminer l'occurrence des malformations développementales et des événements de survie chez les embryons de poisson zèbre exposés à huit concentrations d'acide valproïque (VPA) au cours des cinq premiers jours de vie. Les observations ont été enregistrées quotidiennement pour chaque embryon. La structure et les paramètres du modèle sont déterminés statistiquement par approche bayésienne. La probabilité du taux d'éclosion varie en fonction du temps et nous avons comparé cinq modèles différents pour modéliser sa dépendance au temps. Une fonction constante par partie du temps décrit de manière adéquate les données d'éclosion. Les autres taux de transition sont conditionnés par la concentration embryonnaire de VPA, obtenue avec un modèle pharmacocinétique basé sur la physiologie. Le VPA a principalement impacté le taux de probabilité de malformation dans les embryons éclos et non éclos. Les taux de probabilité de réversion des malformations ont été abaissés par l'VPA. La mortalité directe était faible aux concentrations testées, mais augmentait linéairement avec la concentration interne. Le modèle utilise pleinement les données et donne une analyse plus fine des effets tératogènes du VPA chez le poisson zèbre que l'approche prescrite par l'OCDE. Nous discutons de l'utilisation du modèle pour obtenir des valeurs de référence toxicologiques adaptées à l'extrapolation inter-espèces. Une conclusion générale est que les modèles multi-états complexes peuvent être efficacement évalués numériquement.

Multistate Models of Developmental Toxicity: Application to Valproic Acid-Induced Malformations in the Zebrafish Embryo[‡]

Ségolène Siméon[†], Rémy Beaudouin[~], Katharina Brotzmann[‡], Thomas Braunbeck[‡], Frédéric Y. Bois^{*§}.

[†] INERIS, METO unit, Parc ALATA BP2, Verneuil en Halatte, France.

[~] UMR-I 02 SEBIO, INERIS, Parc ALATA BP2, Verneuil en Halatte, France.

[‡] University of Heidelberg, Aquatic Ecology and Toxicology, Centre for Organismal Studies (COS), Im Neuenheimer Feld 504, D-69120 Heidelberg, Germany.

[§] CERTARA UK Limited, Simcyp Division, Level 2-Acero, 1 Concourse Way, Sheffield, S1 2BJ, United Kingdom. Email: * frederic.bois@certara.com

Abstract

For the determination of acute toxicity of chemicals in zebrafish (*Danio rerio*) embryos, the OECD test guideline 236 stipulates a dose-response analysis of four lethal core endpoints and a quantitative characterization of abnormalities including their time-dependency. Routinely, the data are analysed at the different observation times separately. However, observations at a given time strongly depend on the previous effects and should be analysed jointly with them. To solve that problem, we developed multistate models for occurrence of developmental malformations and life events in zebrafish embryos exposed to eight concentrations of valproic acid (VPA) the first five days of life. Observations were recorded daily per embryo. We statistically infer on model structure and parameters using a numerical Bayesian framework. Hatching probability rate changed with time and we compared five forms of its time-dependence. A piecewise constant function of time adequately described the hatching data. The other transition rates were conditioned on the embryo body concentration of VPA, obtained using a physiologically-based pharmacokinetic model. VPA impacted mostly the malformation probability rate in hatched and non-hatched embryos. Malformation reversion probability rates were lowered by VPA. Direct mortality was low at the concentrations tested, but increased linearly with internal concentration. The model makes full use of data and gives a finer grain analysis of the teratogenic effects of VPA in zebrafish than the OECD-prescribed approach. We discuss the use the model for obtaining toxicological reference values suitable for inter-species extrapolation. A general result is that complex multistate models can be efficiently evaluated numerically.

Keywords

Multistate model, Zebrafish embryo, Malformations, Survival analysis, Dose-response analysis.

[‡] Toxicology and Applied Pharmacology.

1. Introduction

For human risk assessment, many *in vitro* and *in vivo* tools have been developed to improve the extrapolation of toxic effects from animals or non-animal systems to humans. Given its multiple advantages such as simple husbandry, transparency, small size, rapid development, documented genetic homologies to humans and animal welfare considerations, the zebrafish (*Danio rerio*) embryo has become an attractive species for toxicological and pharmacological testing (Hill et al., 2005; Kantae et al., 2016; Larisch et al., 2017; MacRae and Peterson, 2015; Sipes et al., 2011). For developmental toxicity testing in the zebrafish embryo, the OECD test guideline (TG) 236 on the Fish Embryo Acute Toxicity Test (Busquet et al., 2014; Embry et al., 2010; OECD, 2013) prescribes to record counts of four core endpoints (coagulation, lack of heart beat, somite formation and tail detachment) in addition to reporting malformations at various times following exposure of the fish to five test concentrations. The same guideline recommends performing a separate dose-response analysis for the determination of the lowest observed effect (LOEC) as well as LC_{50} and LC_{100} . Only recently, full dose-response curves for all observations have been recommended as an advanced approach for a more detailed data analysis (Brotzmann et al., submitted).

The standard stipulated by OECD TG 236 might also be improved by the following quantitative and qualitative principles: First, translation of toxicity assessment from zebrafish to humans should account for differences in pharmacokinetics between the two species (Quignot et al., 2014); this has been addressed in a previous article on physiologically-based pharmacokinetic (PBPK) modeling for the zebrafish embryo (Siméon et al., 2020). Second, inter-species extrapolation might be improved by a more in-depth mechanistic analysis of the animal model, which would also contribute to a refinement of the test method in line with the 3Rs principle (Russell and Burch, 1959). Although the design of mechanistic models for developmental toxicity is difficult (Battistoni et al., 2019), at least the general features of the malformation process should be accounted for. In quantitative terms, there are two major problems with a time-stratified statistical analysis: Events recorded at a given time are at least partly dependent on previously manifested effects; as a consequence, malformations occurrences are not independent. For example: embryos may die before completion of heart development; since dead embryos cannot subsequently develop heart malformation, the prevalence of cardiac malformation at lethal test concentrations seems to decrease – a typical intricacy of competing risks. Furthermore, stratifying the data by times of observation reduces

the number of data points per analysis, potentially making their analysis impossible or at least reducing their power considerably.

The standard analysis of such time-of-occurrence data uses survival models (Bradburn et al., 2003a, 2003b; Clark et al., 2003a, 2003b; Jiang and Fine, 2007; Lee and Go, 1997), which are commonly used in clinical studies or failure-time analyses (Gould et al., 2015; Karmen et al., 2019). Multistate Markov models (Andersen et al., 2012; Beyersmann et al., 2012; Jackson, 2011) are extensions of survival models and are able to deal with competing risks. The present communication illustrates the application of multistate models to the toxicity analysis of zebrafish embryo malformation data following exposure to a model compound: valproic acid (VPA). VPA is a notorious teratogenic antiepileptic and thymoregulator widely used in humans, even though its mechanism of action is not fully understood (Chuang et al., 2012; Fathe et al., 2014; Phiel et al., 2001). The final model includes five physiological states the zebrafish embryos may be in: “normal” (*i.e.*, pre-hatched), “hatched”, “non-hatched with effects” (*i.e.*, malformed), “hatched with effects” and “coagulated” (*i.e.*, dead). The probabilities of transitions between states may depend on time (*e.g.*, for hatching) or on internal chemical exposure concentration (*e.g.*, for the transitions to malformed or dead states). Data collected at all times and all exposures levels were jointly analyzed, and an innovative route for solving the associated equations is proposed. Additionally, we computed internal VPA concentrations as a function of time and water-concentration using a zebrafish embryo physiologically-based pharmacokinetic (PBPK) model previously developed (Siméon et al., 2020). Combining PBPK and multistate effect models allows to establish concentration-time-response relationships, which can be used for the establishment of reference values.

2. Background on multistate models

Survival analysis models the statistical distribution of the time of occurrence of events of interest (*e.g.*, death), after a fixed starting time, using data from several subjects or observation units (patients, zebrafish embryos, *etc.*) (Bradburn et al., 2003a). The simplest survival model (Figure 1) considers only two states, *e.g.*, “alive” and “dead”, with no possible reversion from the second to the first. The deaths observed correspond to transitions between states 1 and 2, characterized by a hazard rate (also called transition intensity) $q_{1,2}$. This transition intensity is of fundamental interest, because it can link instantaneous exposure to drugs or toxicants and observed death times in exposed individuals.

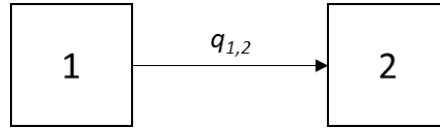


Figure 1. Schematic representation of a basic two-state survival model.

Survival analysis can be generalized to multistate Markov models, which describe transitions, eventually reversible, between more than two states and can be used to obtain the distribution of residence time in the various states (Andersen et al., 2012; Beyersmann et al., 2012; Welton and Ades, 2005). If several states are accessible from another one (Figure 2), we have a competing risks process.

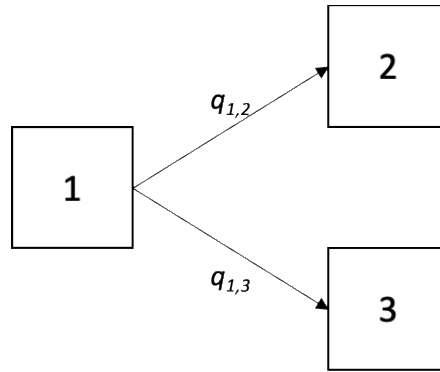


Figure 2. Schematic representation of a simple competing risk model. After a transition from states 1 to 2, state 3 cannot be reached.

Therefore, multistate Markov models form a general class of competing risk models, and can consider many intermediate events (Farewell and Tom, 2014; Putter et al., 2007) (see for example Figure 3).

In these models, a set of transition intensities $q_{i,j}(t, z(t))$ determines the risk of moving from state i to state j , as a function of time and individual-specific variables $z(t)$, which can themselves be time-dependent (Jackson, 2011). An important class of variables $z(t)$ are exposures to hazards, measured for example by the body or organ concentration of a toxicant.

Assume we start from state i among a set of states X , and call $p_{i,j}(t)$ the transition probability from state i to state j during time length t . By definition:

$$p_{i,j}(t) = P\{X(t) = j | X(0) = i\} \quad (1)$$

Determining these transition probabilities is important, because they condition the likelihood of observing transitions between states as time passes (*i.e.*, the data). To calibrate a

multistate model from data, we need to link variables $z(t)$, transition intensities $q_{ij}(t, z(t))$, transition probabilities $p_{ij}(t)$ and observations. A system of ordinary differential equations (ODE), called forward Kolmogorov differential equations, can be written to link transition intensities and probabilities (Brinks and Hoyer, 2018; Fisz, 1976). This system of equations can be expressed quite simply with a matrix notation:

$$\frac{\partial \mathbf{P}(t)}{\partial t} = \mathbf{P}(t) \mathbf{Q}(t) \quad (2)$$

The transition probabilities $p_{ij}(t)$ are the elements of the probability matrix $\mathbf{P}(t)$, and the corresponding transition intensities $q_{ij}(t, z(t))$ are the elements of the transition matrix $\mathbf{Q}(t)$. The ODE system is obtained by element-wise multiplication of the rows of matrix $\mathbf{P}(t)$ by the columns of matrix $\mathbf{Q}(t)$. The probability matrix $\mathbf{P}(t)$ we need is the solution of the above system of differential equations (Jackson, 2011) with suitable initial conditions (*i.e.*, $\mathbf{P}(0) = \mathbf{I}$, the identity matrix).

The final link between transition probabilities and data likelihood depends on the observational process, *i.e.* on study design and type of data collected (Jackson, 2011). In an intermittently-observed process, like the OECD 236 test, the data is a set of n observation times $\{t_{1,m}, \dots, t_{n,m}\}$ and states $\{s(t_{1,m}), \dots, s(t_{n,m})\}$ in which individual m was at these times. This is also called time-censored “panel” data, because the exact time of occurrence of the malformation is not known with precision. The likelihood of observing a pair of successive states $(s(t_{k,m}), s(t_{k+1,m}))$ at times $(t_{k,m}, t_{k+1,m})$ is simply, by definition, the probability of going from state $s(t_{k,m})$, to state $s(t_{k+1,m})$ during the length of time $t_{k+1,m} - t_{k,m}$:

$$L_{k,m} = p_{s(t_{k,m})s(t_{k+1,m})}(t_{k+1,m} - t_{k,m}) \quad (3)$$

The full likelihood of the data, L , is simply the product of each $L_{k,m}$ over all observed transitions and individuals. This essentially amounts to assuming a multinomial distribution of transition counts. Observations times do *not* need be equally spaced for eq. 3 to apply.

Two simplifications of the above general model can be made:

- Time-homogeneous multistate Markov models consider that transition intensities only depend on the starting state and their parameters do not depend on time (Putter et al., 2007; Zare et al., 2014):

$$q_{i,j}(t, z(t)) = q_{i,j} \quad (4)$$

In this case, the ODE system is linear and can be solved analytically using the so-called matrix exponentials (Jackson, 2011; Jones et al., 2017), which can be quite computationally expensive (Tsiros et al., 2019).

- Time-homogeneous semi-Markov models consider that transition intensities depend only on the starting state i and on the entry time t_j into the next state j (Meira-Machado et al., 2009):

$$q_{i,j}(t, z(t)) = q_{i,j}(t - t_j) \quad (5)$$

The model can also be extended to hidden Markov processes if, for example, the exact states may be identified wrongly or indirectly by the observations. We will not consider such hidden Markov models here because the definition and observation of malformations are quite crisp in the zebrafish embryo. However, we are interested in general non-time-homogeneous models because transition intensities may change continuously with zebrafish embryo age. In particular, embryos grow and their internal concentrations of toxicants changes continuously with time (Siméon et al., 2020). In that case, the matrix exponential solution does not apply, and numerical integration of the ODE system is the best solution, which we use.

3. Materials and methods

3.1. Zebrafish embryo malformation data

Spawning groups were obtained from a colony of adult wild-type zebrafish (*Danio rerio*) of the ‘Westaquarium strain’ maintained at the fish facilities of the Aquatic Ecology and Toxicology Group at the University of Heidelberg (licensed under no. 35-9185.64/BH). All experiments were based on OECD TG 236 and associated work (Braunbeck et al., 2015; Braunbeck and Lammer, 2006; Embry et al., 2010; Lammer et al., 2009; OECD, 2013); however, as an adjustment, for better observation of effects, exposure of the embryos was extended to 120 hours post-fertilization (hpf). According to Strähle et al. (2012), an extension of the exposure period to 120 hpf is possible without violation of the current EU animal welfare legislation (EU, 2010) and is recommended in cases of inconclusive observations after 96 hpf.

For each test, embryos were immersed in the test solutions at the 16 cell-stage at the latest (≤ 90 min; before cleavage of blastodisc). To start exposures with minimum delay, twice the number of eggs eventually needed per treatment group were picked from the same batch of eggs and transferred into 100 ml crystallization dishes with the test concentrations or negative (artificial water according to ISO 7346-3) (OECD, 2013) and positive controls (24.7 μ M of 3,4-

dichloroaniline). At 3 hpf at the latest, viable eggs were selected for normal development under the stereomicroscope (≥ 30 -fold magnification) and transferred to a final volume of 1 ml into 24-well plates (one embryo per well). To account for potential adsorption of the test solutions to the plastic walls of the wells, the plates had been pre-exposed to the test solutions for 24 h. Test solutions were replaced at 24, 48, 72 and 96 hpf (without changing the well plates). Embryos were not dechorionated and hatched on their own at approximately 72 hpf. VPA was tested in three independent runs at seven test concentrations ranging from 6.25 to 400 μM regularly spaced by a factor of 2. The control group comprised a total of 120 embryos, while VPA-exposed groups consisted of 60 embryos per test concentration.

Prior to medium replacement and upon completion of the test at 120 hpf, embryos were analyzed for macroscopically discernable alterations including the four morphological core endpoints listed by OECD TG 236 (OECD, 2013) (coagulation, lack of heart-beat, somite formation and tail detachment) as well as any additional sublethal observation including, e.g., scoliosis/lordosis, eye deformation, loss of pigmentation, various types of edemata and general skeletal deformations (Braunbeck et al., 2005; Hollert et al., 2003; Nagel, 2002). Up to 55 developmental malformations could be recorded including a semi-quantitative graduation of severity for at least part of the adverse effects, e.g. heart-beat reduced, severely reduced or missing. All the effects recorded, and their corresponding codes are summarized in Supplemental Material Table S1.

For documentation, morphological alterations were recorded with a Zeiss Axio Cam ICc1 camera mounted on a Zeiss Olympus CKX41 microscope (Carl Zeiss, Oberkochen, Germany) and analyzed using the Zeiss imaging program Zen lite 2011.

Since the addition of buffer may affect the development of the embryos, pH of the artificial water was not buffered, but measured independently at the various test concentrations of VPA (Supplemental Material Table S2). Water pH decreased with VPA concentration, which was expected and taken into account for internal concentration estimation by the PBPK model (see below). However, since the pH dropped below 6.0 for the highest VPA test concentration (800 μM), leading to rapid death of the embryos within a couple of hours by a mechanism completely independent of malformations and subsequent lethality, data obtained at 800 μM were excluded from this analysis.

3.2. Zebrafish embryo malformation multistate model

3.2.1. Model structure

The proposed multistate model of malformations in the zebrafish embryo includes five states, as shown in Figure 3: “normal”, N, *i.e.*, pre-hatched; “hatched”, H; “non-hatched with effects”, E, *i.e.*, malformed; “hatched with effects”, EH; “coagulated”, C, *i.e.*, dead. We aggregated the various malformations into non-hatched with effects and hatched with effects for simplicity.

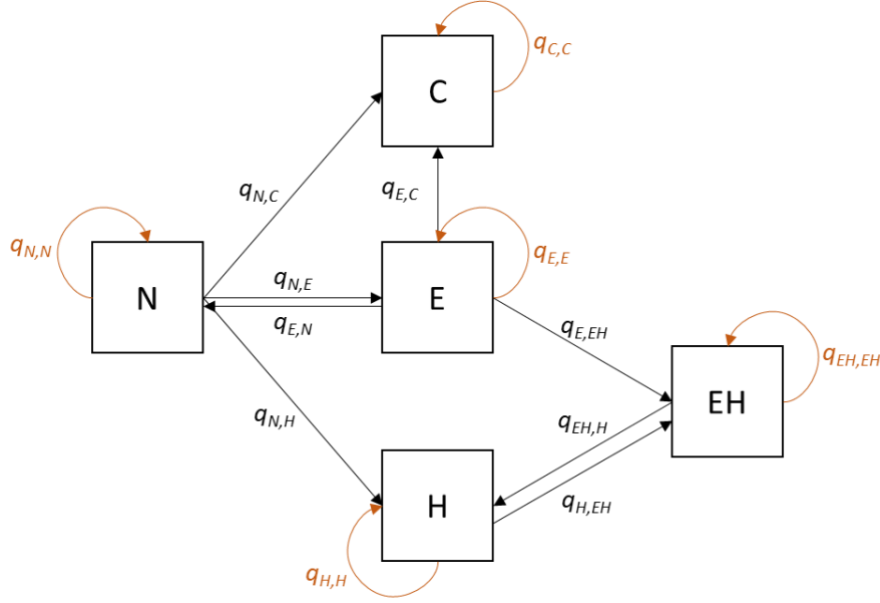


Figure 3. Schematic representation of our five-state competing risk model for the zebrafish embryo: “normal” (N), “hatched” (H), “non-hatched with effects” (E), “coagulated” (C), and “hatched with effects” (EH). The transition intensities from one state to itself (instantaneous probability of staying in a state) have been made explicit.

The coagulated state is the only absorbing state (state from which you cannot return). Back transitions are possible between the normal and non-hatched with effects states, and between the hatched and hatched with effects states, since we had observations of such transitions.

The elements of the associated transition probability matrix $\mathbf{P}(t)$ are the transition probability $p_{i,j}(t)$. According to our model, the transition probabilities $P_{H,N}(t)$, $P_{C,N}(t)$, $P_{EH,N}(t)$, $P_{C,H}(t)$, $P_{H,E}(t)$, $P_{C,E}(t)$, $P_{EH,E}(t)$, $P_{H,C}(t)$, $P_{EH,C}(t)$, and $P_{C,EH}(t)$ will always be null because there is no path between the first and the second state. Hence, there is no need to compute a differential equation for them. Note however, for example, that while there is no direct

transition from the normal to the hatched with effects states, this transition is still possible via the non-hatched with effects state. The transition probability $P_{C,C}(t)$ is equal to 1 because C is absorbing state. Therefore, $\mathbf{P}(t)$ can be written as follow:

$$\mathbf{P}(t) = \begin{pmatrix} p_{N,N}(t) & p_{N,H}(t) & p_{N,E}(t) & p_{N,C}(t) & p_{N,EH}(t) \\ 0 & p_{H,H}(t) & 0 & 0 & p_{H,EH}(t) \\ p_{E,N}(t) & p_{E,H}(t) & p_{E,E}(t) & p_{E,C}(t) & p_{E,EH}(t) \\ 0 & 0 & 0 & 1 & 0 \\ 0 & p_{EH,H}(t) & 0 & 0 & p_{EH,EH}(t) \end{pmatrix} \quad (6)$$

For the associated transition matrix, $\mathbf{Q}(t)$, the row sums have to be equal to 0. Therefore, the diagonal transition intensities $q_{N,N}(t)$, $q_{H,H}(t)$, $q_{E,E}(t)$, $q_{C,C}(t)$, and $q_{EH,EH}(t)$ must be respectively:

$$q_{N,N}(t) = -(q_{N,H}(t) + q_{N,E}(t) + q_{N,C}(t)) \quad (7)$$

$$q_{H,H}(t) = -q_{H,EH}(t) \quad (8)$$

$$q_{E,E}(t) = -(q_{E,N}(t) + q_{E,C}(t) + q_{E,EH}(t)) \quad (9)$$

$$q_{C,C}(t) = 0 \quad (10)$$

$$q_{EH,EH}(t) = -q_{EH,H}(t) \quad (11)$$

The transition matrix $\mathbf{Q}(t)$ is therefore:

$$\mathbf{Q}(t) = \begin{pmatrix} -q_{N,H}(t) - q_{N,E}(t) - q_{N,C}(t) & q_{N,H}(t) & q_{N,E}(t) & q_{N,C}(t) & 0 \\ 0 & -q_{H,EH}(t) & 0 & 0 & q_{H,EH}(t) \\ q_{E,N}(t) & 0 & -q_{E,N}(t) - q_{E,C}(t) - q_{E,EH}(t) & q_{E,C}(t) & q_{E,EH}(t) \\ 0 & 0 & 0 & 0 & 0 \\ 0 & q_{EH,H}(t) & 0 & 0 & -q_{EH,H}(t) \end{pmatrix} \quad (12)$$

The model ODE system is obtained by a simple matrix multiplication of $\mathbf{P}(t)$ by $\mathbf{Q}(t)$:

$$\frac{\partial p_{N,N}(t)}{\partial t} = -p_{N,N}(t) \times (q_{N,H}(t) + q_{N,E}(t) + q_{N,C}(t)) + p_{N,E}(t) \times q_{E,N}(t) \quad (13)$$

$$\frac{\partial p_{N,H}(t)}{\partial t} = p_{N,N}(t) \times q_{N,H}(t) - p_{N,H}(t) \times q_{H,EH}(t) + p_{N,EH}(t) \times q_{EH,H}(t) \quad (14)$$

$$\frac{\partial p_{N,E}(t)}{\partial t} = p_{N,N}(t) \times q_{N,E}(t) - p_{N,E}(t) \times (q_{E,N}(t) + q_{E,C}(t) + q_{E,EH}(t)) \quad (15)$$

$$\frac{\partial p_{N,C}(t)}{\partial t} = p_{N,N}(t) \times q_{N,C}(t) + p_{N,E}(t) \times q_{E,C}(t) \quad (16)$$

$$\frac{\partial p_{N,EH}(t)}{\partial t} = p_{N,H}(t) \times q_{H,EH}(t) + p_{N,E}(t) \times q_{E,EH}(t) - p_{N,EH}(t) \times q_{EH,H}(t) \quad (17)$$

$$\frac{\partial p_{H,H}(t)}{\partial t} = -p_{H,H}(t) \times q_{H,EH}(t) + p_{H,EH}(t) \times q_{EH,H}(t) \quad (18)$$

$$\frac{\partial p_{H,EH}(t)}{\partial t} = p_{H,H}(t) \times q_{H,EH}(t) - p_{H,EH}(t) \times q_{EH,H}(t) \quad (19)$$

$$\frac{\partial p_{E,N}(t)}{\partial t} = -p_{E,N}(t) \times (q_{N,H}(t) + q_{N,E}(t) + q_{N,C}(t)) + p_{E,E}(t) \times q_{E,N}(t) \quad (20)$$

$$\frac{\partial p_{E,H}(t)}{\partial t} = p_{E,N}(t) \times q_{N,H}(t) - p_{E,H}(t) \times q_{H,EH}(t) + p_{E,EH}(t) \times q_{EH,H}(t) \quad (21)$$

$$\frac{\partial p_{E,E}(t)}{\partial t} = p_{E,N}(t) \times q_{N,E}(t) - p_{E,E}(t) \times (q_{E,N}(t) + q_{E,C}(t) + q_{E,EH}(t)) \quad (22)$$

$$\frac{\partial p_{E,C}(t)}{\partial t} = p_{E,N}(t) \times q_{N,C}(t) + p_{E,E}(t) \times q_{E,C}(t) \quad (23)$$

$$\frac{\partial p_{E,EH}(t)}{\partial t} = p_{E,H}(t) \times q_{H,EH}(t) + p_{E,E}(t) \times q_{E,EH}(t) - p_{E,EH}(t) \times q_{EH,H}(t) \quad (24)$$

$$\frac{\partial p_{EH,H}(t)}{\partial t} = -p_{EH,H}(t) \times q_{H,EH}(t) + p_{EH,EH}(t) \times q_{EH,H}(t) \quad (25)$$

$$\frac{\partial p_{EH,EH}(t)}{\partial t} = p_{EH,H}(t) \times q_{H,EH}(t) - p_{EH,EH}(t) \times q_{EH,H}(t) \quad (26)$$

As mentioned above, the other differentials, *e.g.* for $P_{H,N}(t)$, are null. Similarly, $P_{C,C}(t)$ is always equal to 1 and its differential is null. To check the model equations, we can calculate the probability mass balance, which should always be equal to 5 because the initial condition for $P(t)$ is the identity matrix of order 5:

$$\text{Probability Mass Balance} = \sum_{i=1}^5 \left(\sum_{j=1}^5 P_{i,j}(t) \right) = 5 \quad (27)$$

3.2.2. Transition hatching rate sub-model

The zebrafish embryo hatching rate is clearly non-constant with time: in normal conditions, hatching is never observed before 24 hpf because the embryo is not ready for hatching (Kimmel et al., 1995). We tested five sub-models linking $q_{N,H}(t)$ and time and retained a piecewise constant model with estimated hatching time on the basic of the Akaike information criterion (see Supplemental Material Section 2). In the five-state model (Figure 3), hatching can happen to normal embryos (state N) and malformed embryos (state E). The piecewise constant sub-model was used for both hatching rates in the final model, leading to three parameters to estimate: hatching time t_h , and the values of the intensities $q_{N,H}$ and $q_{E,EH}$ after t_h (before t_h those intensities were assumed to be null).

3.2.3. VPA concentration-dependent transition intensities

To find appropriate functional forms for the links between VPA internal concentration, $C(t)$, and transition intensities, we first estimated the remaining intensities ($q_{E,N}$, $q_{EH,H}$, $q_{N,E}$, $q_{N,C}$, $q_{E,C}$, and $q_{H,EH}$) separately for each exposure group. We then determined visually the needed sub-models: exponential for $q_{E,N}$ and $q_{N,E}$; linear for $q_{EH,H}$, $q_{N,C}$ and $q_{E,C}$; and Hill for $q_{H,EH}$. The corresponding equations are:

$$q_{E,N}(C(t)) = q_{E,N}(0) \times e^{k_{q_{E,N}} \times C(t)} \quad (28)$$

$$q_{N,E}(C(t)) = q_{N,E}(0) \times e^{k_{q_{N,E}} \times C(t)} \quad (29)$$

$$q_{EH,H}(C(t)) = q_{EH,H}(0) + k_{q_{EH,H}} \times C(t) \quad (30)$$

$$q_{N,C}(C(t)) = q_{N,C}(0) + k_{q_{N,C}} \times C(t) \quad (31)$$

$$q_{E,C}(C(t)) = q_{E,C}(0) + k_{q_{E,C}} \times C(t) \quad (32)$$

$$q_{H,EH}(C(t)) = \frac{q_{H,EH_{max}} \times C(t)^n}{EC_{50}^n + C(t)^n} \quad (33)$$

where $q_{E,N}(0)$, $q_{N,E}(0)$, $q_{EH,H}(0)$, $q_{N,C}(0)$, and $q_{E,C}(0)$ the transition intensities at VPA concentration zero; $k_{q_{E,N}}$, $k_{q_{N,E}}$, $k_{q_{EH,H}}$, $k_{q_{N,C}}$, and $k_{q_{E,C}}$ are constants; $q_{H,EH_{max}}$ is the maximal transition intensity value of $q_{H,EH}$, n its Hill exponent, and EC_{50} the concentration for which $q_{H,EH}$ is equal to 50% of its maximal value.

3.3. Coupling to a zebrafish embryo PBPK model

The internal VPA concentration vs. time profiles (whole body excluding yolk) driving the above transition intensities (Eqs. 28 to 33) were computed for the various nominal exposure concentrations tested using our previously published zebrafish embryo PBPK model (Siméon et al., 2020). The decrease in water pH with increasing VPA test concentration was taken into account. The pre-computed internal concentration vs. time profiles were read in the start of the multistate calculations and interpolated linearly as needed during integration of the ODE equations. Figure S1, in Section 1 of the Supplemental Material shows the value of internal concentration as a function of nominal concentration and time (on the Figure, discontinuities are due to the daily medium changes).

3.4. Statistical inference on multistate model parameters

The multistate parameters of the joint PBPK-multistate model were calibrated using Markov chain Monte Carlo (MCMC) simulations in a Bayesian statistical framework (Bernillon and Bois, 2000; Bois, 2012; Smith and Roberts, 1993). MCMC methods can draw samples from the joint multi-dimensional posterior distribution of the model parameters, given a statistical model, prior parameter distributions, and data for which a likelihood function can be computed (Bois, 2009).

We used the following statistical model: For most parameters ($q_{E,N}(0)$, $q_{N,E}(0)$, $q_{EH,H}(0)$, $q_{N,C}(0)$, $q_{E,C}(0)$, $kq_{E,N}$, $kq_{N,E}$, $kq_{EH,H}$, $kq_{N,C}$, $kq_{E,C}$, $q_{H,EHmax}$, n , and EC_{50}) we estimated one common value given the data from all experimental exposure groups. The three hatching parameters (t_h , $q_{N,H}$, and $q_{E,EH}$) were estimated at the same time, but in a multi-level framework (Bois, 2012): As indicated in the experimental protocol, because the embryos of each exposure may not have been exactly at the same morulation stage at the time of selection, and due to biological variability, average hatching time varied randomly between exposure groups. Therefore, at the higher level of the hierarchy, we have three mean hatching parameters, and for individual exposure groups we assume that group-specific values are distributed randomly around the common mean:

$$\boldsymbol{\varphi}_i \sim \text{Lognormal}(\boldsymbol{\varphi}, \boldsymbol{\sigma}) \quad (34)$$

where $\boldsymbol{\varphi}$ is the geometric mean in log-space of either t_h , $q_{N,H}$, or $q_{E,EH}$, and $\boldsymbol{\sigma}$ is the corresponding standard deviation in log-space. The lognormal distribution for t_h was truncated to fall between 0 and 5 (days). So, for hatching, three mean parameters and eight triplets of group-specific parameters were estimated.

Non-informative uniform priors were used for most transition intensities' parameters. Parameters $q_{EH,H}(0)$ and $kq_{EH,H}$ were sampled with correlated constraints to respect the fact that $q_{EH,H}$ must be positive. Their sampling was done *via* two auxiliary variables, $uq_{EH,H}(0)$ and $vkq_{EH,H}$, which were the parameters actually estimated. Parameter $q_{EH,H}(0)$ was then assigned value $-uq_{EH,H}(0)$, and $kq_{EH,H}$ was assigned the value $vkq_{EH,H} / 0.02$. This ensures that $q_{EH,H}$ is always positive. For the SDs of the dose-group-specific hatching parameters, very wide truncated normal distributions were used (Table 1).

The data were censored by interval. The data likelihood was therefore computed as the product of terms given by Eq. 3 over all transitions and individual embryos.

Two simulated Markov chains of 30000 iterations each were generated, starting from a set of parameter values randomly sampled from their priors. Convergence of the last 15000 iterations of two chains was assessed using Gelman and Rubin's Rhat convergence criterion (Gelman and Rubin, 1992).

Table 1. Prior parameters distribution used for MCMC calibration of the five-state model with embryo transition count data.

Parameter	Prior distribution ^a
$q_{E,N}(0)$	$U(0, 2)$
$k_{q_{E,N}}$	$U(-6 \times 10^2, 0)$
$u_{q_{EH,H}}(0)$	$U(-4, 0)$
$v_{k_{q_{EH,H}}}$	$U(u_{q_{EH,H}}(0), 0)$ ^b
$q_{N,E}(0)$	$U(0, 0.06)$
$k_{q_{N,E}}$	$U(2 \times 10^2, 4.5 \times 10^2)$
$q_{N,C}(0)$	$U(0, 0.01)$
$k_{q_{N,C}}$	$U(0, 10)$
$q_{E,C}(0)$	$U(0, 0.1)$
$k_{q_{E,C}}$	$U(0, 10)$
$q_{H,EHmax}$	$U(0, 40)$
n	$U(0, 15)$
EC_{50}	$U(0, 0.01)$
$q_{N,H}$ ^c	$U(0, 20)$
$q_{E,EH}$ ^c	$U(0, 10)$
t_h ^c	$U(1, 5)$
$\sigma_{q_{N,H}}$	$N_t(1.1, 0.5, 1.01, 10)$
$\sigma_{q_{E,EH}}$	$N_t(1.1, 0.5, 1.01, 10)$
σ_{th}	$N_t(1.1, 0.5, 1.01, 10)$

^a Distribution shapes: $U(a, b)$ for uniform between a and b ; $N_t(m, s, a, b)$ for truncated normal with mean m , SD s , and bounds a and b .

^b Correlated sampling.

^c Prior distribution of the inter-group mean.

3.5. Calculation of predicted embryos counts

The number of zebrafish embryos in state j at time t , $n_j(t)$, is conditioned by the probability to go from each initial state i to state j , according to the number of zebrafish embryos in initial state i at starting time t_0 , $n_i(t_0)$, and is defined by:

$$n_j(t) = \sum (n_i(t_0) \times P_{i,j}(t)) \quad (35)$$

3.6. Calculation of reference toxicity values

Internal concentrations values leading to 10% effects (internal EC_{10}) on intensities were simply computed by mathematical inversion of Equations 28 to 33 (see Supplemental Material, section 4). This was done for all the parameter values sampled by MCMC to obtain distributions of EC_{10} values reflecting uncertainty. External concentrations values leading to 10% effects (external EC_{10}) on probabilities were computed by running the joint PBPK-multistate model for a grid of VPA water concentrations. Estimated probabilities were recorded for a dense set of times, leading to sets of three-dimensional exposure-time-response relationships. Contour at z-values 0.1 or 0.9 (depending on whether P increases or decreases with VPA exposure, respectively) were obtained with the “*contour*” function of the R statistical package for the maximum posterior parameter set and 100 random parameter sets, to estimate uncertainty.

3.7. Software

The dynamic model simulations and MCMC calibrations were performed with GNU MCSIM version 6.1.0 (<https://www.gnu.org/software/mcsim>) (Bois, 2009). We used the R software version 3.5.2 for plotting and miscellaneous statistical analyses (R Development Core Team, 2013). The model code and the MCMC specification file are shown in Supplemental Material, Section 6.

4. Results

4.1. Model calibration

4.1.1. Model fit to the data

Each MCMC chain of 30000 iterations took 9 hours to complete on a Dell Precision 5530 laptop computer (with Intel Core i7-8850H CPU clocked at 2.6GHz) using the Linux subsystem. Figure 4 shows all observed transition counts plotted against the corresponding maximum posterior (best) model predictions. Note that the predictions are expected values, rather than integer counts, so there is a layer of randomness in the data that is not taken into account in this plot. Note also that the graph uses logarithmic scales for readability. However, since null values (mostly for observations) cannot be shown on such a plot, we assigned them an arbitrary value of 0.1, only for plotting purposes; some of them fall far away from the identity line because of that artifact. The model tends to under-predict the data, but most of the points fall in the two-fold error interval. Table S4 (Supplemental material) gives all predicted and observed values. If we disregard the null values, the most outlying points are from the control

group (three transitions observed from normal to hatched with effects on day 4, and two embryos remaining in the effect state on day 5) or correspond to observations of one or two transitions in the exposed groups while the model predicts a low probability for those events.

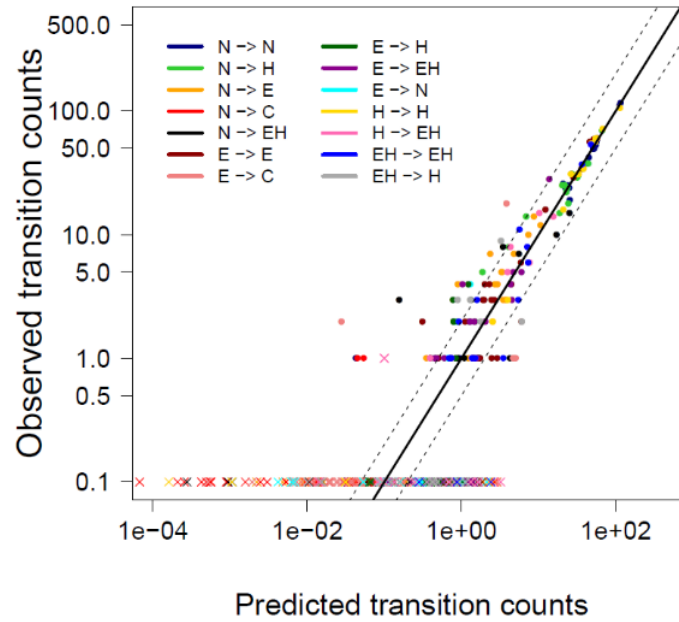


Figure 4. Observed versus best model-predicted transition counts. The black line is the perfect fit. The dashed lines give the two-fold-error band. The crosses are null observations or predictions assigned an arbitrary value of 0.1 to appear on the graph.

Figure 5 shows the detailed fit of the coupled PPBK-multistate model to the number of transitions between states, for the control group and the highest exposure concentration. The model predictions for all concentrations are presented in the Figure S7 of the Supplemental Material. As VPA exposure concentration increases, we observe a decrease in transitions to the hatched state, and an increase in transitions to the malformed, hatched malformed, and dead states. Yet, most exposure concentrations have little effect. Time has a strong effect on transitions between states because hatching intensities are not constant in time (for) and because of competing risk and depletion effects. For example, in the control group, the count of transitions from normal to hatched states is high at the end of day 3, because $q_{N,H}(t)$ become non-null, but it decreases at the end of days 4 and 5, simply because there are no more embryos left to hatch. At the highest VPA concentration, the count of transitions from normal to malformed is highest at the end of day 1 because of VPA, but then decreases because of depletion of the normal fish. At the same concentration, the count of malformed to malformed transitions (*i.e.*, the count of fish staying in the malformed state) first increases and then decreases, because malformed embryos start either hatching or dying. The predicted counts of

embryos malformed, hatched with effects, or dead is shown in Figure 6 as a function of the time for each exposure concentration (the predicted counts for all states are shown in Figure S8 of the Supplemental Material). The number of embryos hatched with effects or dead increases with the concentration and the time of exposure, but the number of malformed embryos decreases with time and would not be well nor coherently modelled by standard sigmoidal dose-response curves. Note that there is less information in those count data than in the transition data (less data points for the same experiments), and that is why the model is fitted to the transition data.

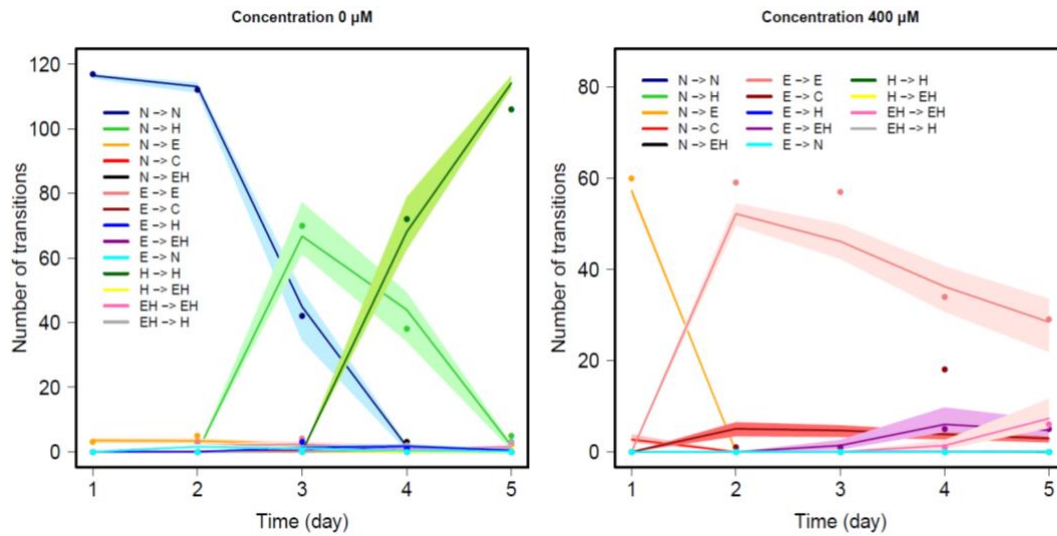


Figure 5. Observed (dots) and predicted (lines) number of transitions between states normal (N), hatched (H), malformed (E), hatched with effects (EH) and dead (C), as a function of time for the exposure concentrations 0 and 400 μM . Hatching rate is modeled by a piecewise constant intensity with estimated hatching time. The colored bands give the predictions' 95% confidence intervals.

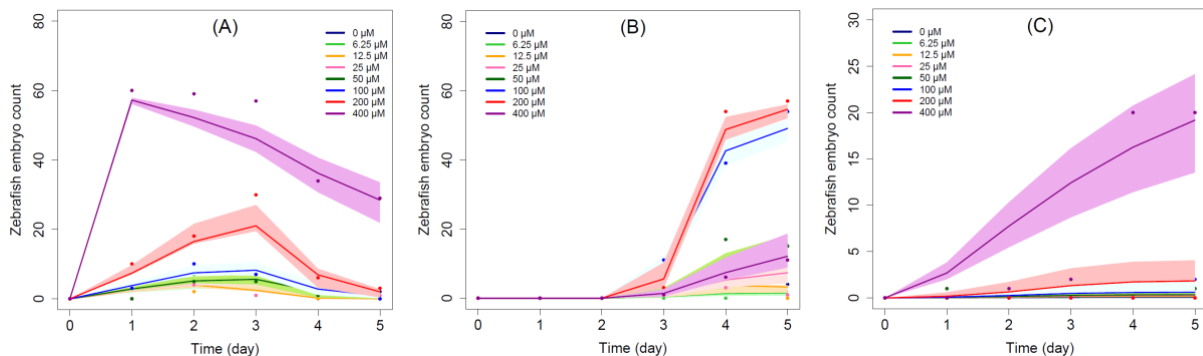


Figure 6. Counts of embryos in the malformed (A), hatched with effects (B) and dead (C) states, as a function of time and VPA concentrations. The dots are the data, the lines are model predictions. The colored bands give the predictions' 95% confidence intervals.

4.1.2. Transition intensities' estimates

The marginal posterior parameters distributions obtained by MCMC calibration with the embryo transition count data are summarized in Table 2. The parameters were estimated with a precision ranging from about 1.4% (for t_h) to 89% ($k_{qN,C}$). A parameter posterior precision depends largely on the number of observed transitions that are conditioned by that parameter. The posterior summaries of the concentration-group-specific hatching parameter estimates are given in Table S5 of the Supplemental Material; there was no obvious correlation of those with concentration, but the hatching intensity is twice as low in malformed embryos as in normal embryos. Note, however, that MCMC sampling yields parameter vectors (sets) out of the joint posterior distribution of all parameters calibrated, that means that correlations between parameter estimates can be viewed, summarized and taken into account in predictive simulations. The correlations between parameters of the same intensity sub-models are shown in Figure S6 of the Supplemental material.

Table 2. Mean, standard deviation (SD), 95% confidence intervals (IC95%) and maximum posterior (MP) value of the MCMC-sampled multistate model parameters.

Parameter	Mean $\pm$ SD	IC 95%	MP
$q_{E,N}(0)$	0.75 ± 0.21	[0.40; 1.2]	0.61
$k_{qE,N}$	-280 ± 97.1	[-479; -103]	-190
$uq_{EH,H}(0)$	-0.53 ± 0.15	[-0.86; -0.29]	-0.45
$vk_{qEH,H}$	-0.33 ± 0.182	[-0.70; -0.023]	-0.19
$q_{N,E}(0)$	0.042 ± 0.0059	[0.032; 0.055]	0.040
$k_{qN,E}$	371 ± 23	[318; 404]	355
$q_{N,C}(0)$	0.0014 ± 0.0012	$[5.9 \times 10^{-5}; 0.0043]$	7.1×10^{-5}
$k_{qN,C}$	0.78 ± 0.70	[0.027; 2.7]	0.55
$q_{E,C}(0)$	0.017 ± 0.014	[0.00065; 0.050]	0.0028
$k_{qE,C}$	3.7 ± 1.2	[1.3; 6.1]	4.5
$q_{H,EHmax}$	24 ± 8.6	[9.2; 39]	23
n	5.8 ± 0.62	[4.6; 7.1]	5.4
EC_{50}	0.0044 ± 0.00056	[0.00345; 0.0056]	0.0046
$q_{N,H}^a$	3.8 ± 0.61	[2.9; 5.2]	3.3
$q_{E,EH}^a$	0.84 ± 0.29	[0.41; 1.5]	0.57
t_h^a	2.8 ± 0.042	[2.7; 2.9]	2.8
$\sigma_{qN,H}$	1.2 ± 0.19	[1.0; 1.7]	1.0
$\sigma_{qE,EH}$	2.0 ± 0.26	[1.6; 2.6]	2.2
σ_{t_h}	1.0 ± 0.015	[1.0; 1.1]	1.0

^a Inter-group mean.

It is challenging to directly interpret the values of Table 2. More telling is Figure 7, which shows the relationships between transition intensities and internal concentration (Eqs. 28 to 33) obtained when using the above posterior parameter estimates. Those are in fact the fundamental dose-response relationships of the model. First, there are large differences between their shape and range of values. The two largest correspond to rates of malformations. Note that the toxicity to hatched embryos starts occurring at lower dose than toxicity to the non-hatched embryos, but saturates above 0.01 M. Next come the suppression of reversions from malformed to normal embryos (hatched or not). Such events are sometimes observed but become even less likely as the internal concentration of VPA increases. Finally, direct mortality is relatively low at the concentrations tested but seems to increase linearly with internal concentration.

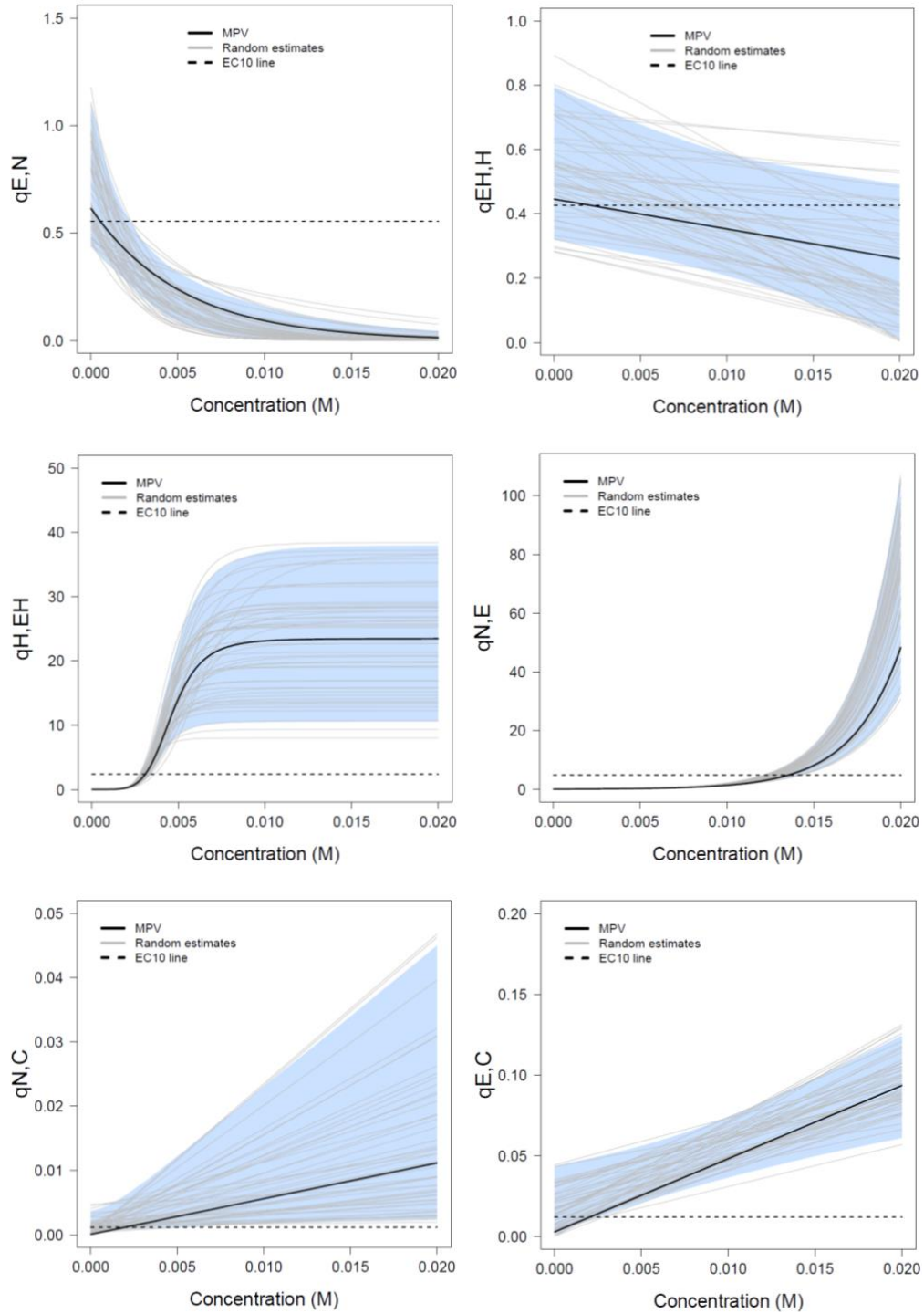


Figure 7. Transition rates as a function of VPA internal concentration. The black line is the best estimate. The grey lines are 50 random estimates. The dashed line corresponds to a 10% change from control (for the best estimate). The blue areas show the 95% confidence interval.

4.2. Calculation of reference toxicity values

Regulatory guidelines on toxicity testing typically recommend to analyze the concentration-response profiles obtained experimentally to identify points of departure from baseline (healthy) state. Such points of departure can, for example, be exposure concentrations (EC_{10}) leading to 10% effect. They can then be transposed to other species for risk assessment. Should such values be sought, they could be derived on the basis of internal or external concentrations and for effects measured by transition intensities (q), or transition probabilities (P). Out of the four possible combinations, we chose to present internal concentrations values leading to 10% effects (EC_{10}) for intensities, and external EC_{10} for probabilities.

The internal EC_{10} values for $q_{EH,H}$, $q_{N,C}$ and $q_{E,C}$, which are linear functions of internal concentration, are always equal to 2×10^{-3} M because uncertainty and slope differences cancel out. The distribution of internal EC_{10} values for $q_{E,N}$, $q_{H,EH}$ and $q_{N,E}$ are shown in Figure 8. The reversion intensity from malformed to normal embryos appears to be the most sensitive endpoint, followed the three linear intensities (reversion from malformed hatched $q_{EH,H}$, direct death $q_{N,C}$ and death of malformed $q_{E,C}$), then by the intensity of malformations in hatched embryos, and finally by the intensity of malformation in non-hatched embryos. Hence, the rate of the primary type of malformations observed appears to be the least sensitive to VPA internal concentration.

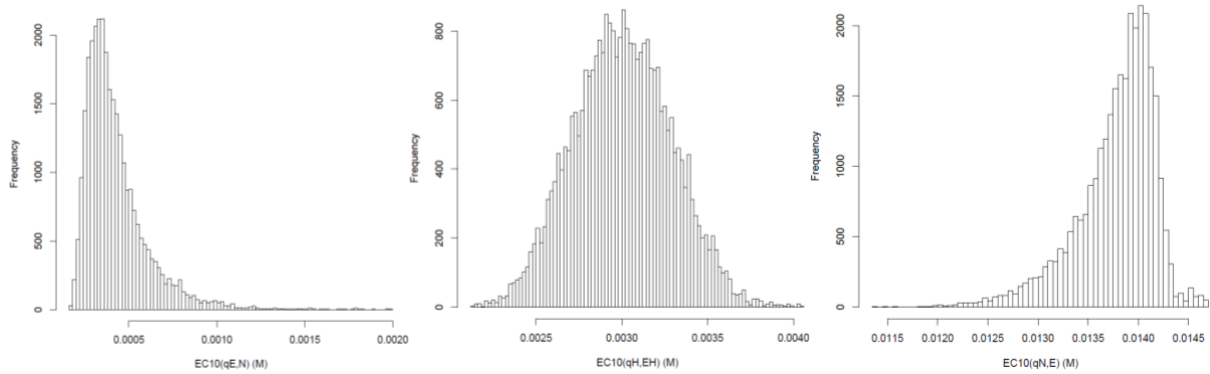


Figure 8. Distribution of EC_{10} values for the transition rates $q_{E,N}$, $q_{H,EH}$, and $q_{N,E}$. Those EC_{10} are defined on a scale of internal embryo concentrations.

Transition intensities are fundamental quantities in the model, but are not directly related to observations of fish counts in the different states, which maybe more relevant for risk assessment. Transition probability conditions more directly fish counts. Internal concentration EC_{10} values for probabilities could be obtained, but it may be more relevant again for risk

assessment, in some context, to obtain EC_{10} values based on external concentration. This can be done, but in those, the EC_{10} depends on exposure concentration and time. Figure 9, top row, shows the 3D water concentration-time-transition probability relationships for transition from normal to malformed, $P_{N,E}$, and normal to dead, $P_{N,C}$. $P_{N,E}$ shows a maximum at mid-exposure concentration and mid-time, because of competition effects with other states (dead, hatched, hatched malformed). $P_{N,C}$ has a less complex shape and increases nonlinearly with time and external concentration, because death is an absorbing state from which the embryo cannot escape. The two bottom panels of Figure 9 show the EC_{10} estimates at different times. These are just contours plots of the 3D surface at 10% of the maximum probability. Obviously, the external concentration leading to 10% increase in transition probability depends on time. For $P_{N,E}$, there is clearly a sensitive period of time just before hatching, when VPA has had time to exert its effect and hatching has not depleted the pool of non-hatched embryos. At that time, the EC_{10} is around 50 μM . For $P_{N,C}$, the longer the lower the EC_{10} but it does not go lower than 200 μM , showing that dead, here also is a less sensitive toxicity endpoint.

Figure 10 shows stradivarius plots comparing external concentration EC_{10} values obtained empirically (see Table S6 in Supplemental material) with the distribution of minimal EC_{10} values based on $P_{N,E}$, and $P_{N,C}$. Those minimal values were obtained by taking the minimum external concentration along each possible contour line. They fall around time 2.8 days for $P_{N,E}$, and as expected, 5 days for $P_{N,C}$. The range of values is comparable, and the most sensitive endpoints found empirically (reduced yolk sac absorption EC_{10} at 20 μM ; Pericardial edemata EC_{10} at 38 μM ; Craniofacial deformation EC_{10} at 52 μM) do fall in the range of the $P_{N,E}$ based EC_{10} 's we found.

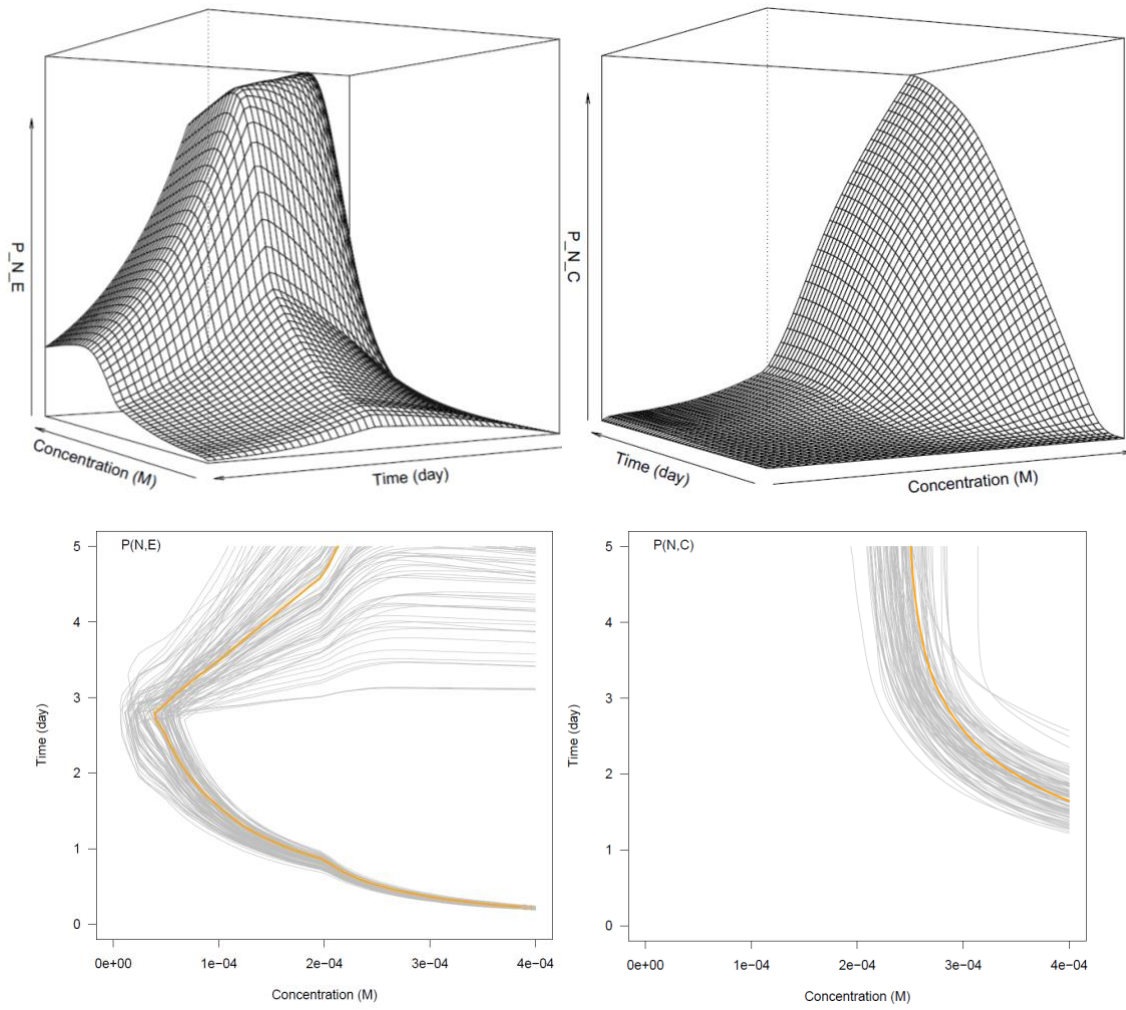


Figure 9. Top row: Probability for an embryo normal at time zero to be in state malformed ($P_{N,E}$, left), or to be dead ($P_{N,C}$, right), as a function of time and VPA water concentration. The parameter best estimates were used to compute the surfaces. Bottom row: Contour lines at 10% probability (EC_{10}) for $P_{N,E}$ (left) and $P_{N,C}$ (right), for the best parameter estimates and 100 random parameter sets.

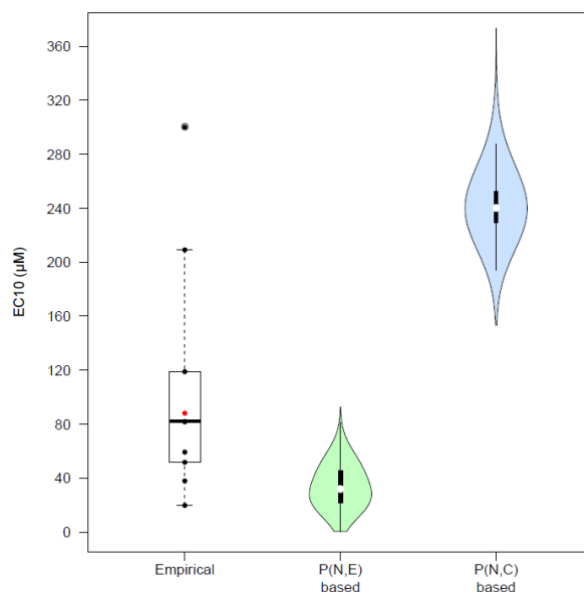


Figure 10. Distributions of the empirically determined external concentration EC_{10} values (the red dot corresponds to death) and of the minimal EC_{10} values obtained for $P_{N,E}$ and $P_{N,C}$ transition probabilities.

5. Discussion

A five-state Markovian model was developed to simulate and predict the occurrence of developmental malformations observed in zebrafish embryo following VPA exposure. The model makes full use of time- and dose-information gathered in the time-course of an OECD 236-compliant zebrafish toxicity test (OECD, 2013). The model takes into account the possibility of reversion of malformations, concurrent transitions between states, time- and exposure-dependent transition intensities. The link to exposure is important, because it offers a semi-mechanistic explanation of the observations, whose correlation pattern is complex and forbidding time by time or end-point by end-point analyses. Using PBPK-predicted internal concentration makes the model even more mechanistic and offers a way to take into account the large pharmacokinetic differences between human and zebrafish embryos, if quantitative extrapolation of the results to humans is sought. However, the model has limitations, as discussed below.

5.1. Model structure

Our model makes large simplification that can partly explain the relative misprediction of low transition count data. For simplicity, the model aggregates in only five states some 55 malformation types which occurred with different intensities and at different times. Differences

in intensities could partly be explained by different organ concentrations of the chemical (which could be predicted by the PBPK, if needed), but also by different organ susceptibilities to malformations. A finer subdivision of states by organs would in turn allow the use of more relevant measures of internal exposure. This would still require calibration of the multistate model parameters, because some chemicals might have an affinity for an organ without causing effects to it, but it would more mechanistic and probably be more predictive after extrapolation to another species. As to timing, it is obvious, for example, that heart malformations will not be observed before heart formation, *etc.*; this is not accounted for in the model. Please also note that some malformation types are in fact varying degree of severity of malformation in the same organ (*e.g.*, for the heart). It is common to use multistate models to describe such progressions (Cannon et al., 2017), and it would be possible to refine our model in such more mechanistic direction. However, one problem arises when partitioning effects into finer states: There is a combinatorial multiplication of pluri-malformation states (*i.e.*, embryos suffering from several malformations at the same time should be in specific states). We do not have a solution to that problem at the moment.

A simplified three-state model was used to find the most appropriate way to describe the time-dependence of hatching intensity. Two hatching sub-models had almost equal AIC values (a piecewise constant model with estimated hatching time and a Hill's model). We preferred the former, and data at a finer time scale (*e.g.*, every few hours) would be needed to refine the model. Note that observations times do not need at all to be equally spaced when using a multistate framework. There was an apparent random effect in hatching intensity between VPA exposure groups and we adopted a hierarchical approach to model it. *A posteriori*, there are not huge or very significant differences in hatching between exposure groups, and only one set of hatching parameters might suffice. While there is no obvious effect of VPA on hatching time or extent, the model estimates separately hatching in normal and malformed embryos, and we found the latter to have a reduced hatching rate. Please note that one type of malformations is “lack of hatch” and could be modeled as a censored piece of data on hatching time.

The shape of the individual sub-models linking internal concentration to transition intensities was determined with a standard two-step approach: First, estimate transition intensities for each exposure group, independently; second, determine empirically the relationship of the estimates with concentration. We did find that all six transitions intensities, not including hatching, discussed above, were VPA concentration-dependent. A more principled method would compare systematically the sub-model fits, or use a non-parametric

model able to describe all of them. However, the computational burden would be much higher, and we would not have enough concentration levels to improve dramatically and validate the intensity sub-models (the four lowest VPA concentrations tested had little effects). Note that at the highest VPA concentration actually tested (800 μM) all embryos had died within the first day, potentially through a direct effect on water pH. We could have modeled this effect by adding a separate term to the death intensity sub-model, very high and effective only at that concentration. But that would be adding one parameter to model one data point, which is quite *ad hoc*. We preferred to focus on malformations and excluded the 800 μM concentration and associated data from our analysis.

There are seven *reverse* transitions going from malformed to normal-looking embryos, and 19 transitions on day 5 from malformed hatched embryos to normal-looking ones. We estimated the corresponding intensities and found them to be negatively affected by VPA. They could have been treated as miss-classified events by using a Hidden Markov extension of the model (Williams et al., 2020), but it seems unlikely that observation errors would be affected by VPA exposure. Furthermore, among the observed malformation (Table S1), some, like edema, are known to be reversible with the embryo's growth and its development its immune system (van der Vaart et al., 2012). Note that an advantage of our global model is that “small” events such as these are identifiable because all time and dose data are analyzed jointly. We would argue that if all 55 malformation types were analyzed, each of them would have low counts, and only such a model would be able to quantify them jointly and adequately.

5.2. Model solving and calibration

We used the transition count data, as is done in multistate modeling, rather than population count data in the various states, as would be typically done in toxicology. Population counts evolve with time and concentration, but they carry less information than transition counts, because they lose information about the trajectory of individual fish through states in time. The standard OECD protocol calls for observations once per day, but indeed, more frequent observations would be more informative. The model takes into account the fact that an observed transition, say from normal to dead, may have occurred through passing (with the day) by the malformed state, without direct observation of it. Making sure that most states are actually observed would bring more precise predictions. Note also that only the three highest doses are strongly informative about effects. Obviously, statisticians always beg for more data, but one application of the model could be to estimate the value of added information and help planning further experiments (Bois et al., 1999; Overstall et al., 2020).

General multistate models can use time-dependent transition intensities. However, in all multistate model applications and software we have seen, the matrix exponential solution is always used and cannot accommodate time-dependent intensities. The traditional solution is to approximate time dependent parameters by piecewise constant functions of time (Williams et al., 2020). However, assuming a piecewise constant process is a strong limitation which makes it difficult and inaccurate to link multistate and continuous pharmacokinetic models, and even more to make inference about the joint pharmacokinetic-pharmacodynamic model. It also increases the number of parameters to estimate. Finally, the matrix exponential solution is clumsy and slow to compute. A much more general solution is to integrate numerically the set of differential equations. Numerical integration imposes no constraints on linearity of the transition intensities or their time independence. Modern integration algorithms are also very fast and precise. Surprisingly, we seem to be the first to have used direct numerical integration of the differential system specifying P in the published literature. Curiously, the statistical community has so far preferred approximating non-linear general models with piecewise-linear models, despite the ensuing lack of flexibility and increased number of parameters to estimate. We compared the solution of the same multistate model using the well-known *msm* R package (Jackson, 2011), which uses in this case the analytical solution (much faster than matrix exponentials) and maximum-likelihood, and our integration method in a Bayesian framework, using *GNU MCSim*. We obtain the same or more precise estimates (see Supplemental material, section 5 and Figure S9), but *GNU MCSim* took 20 minutes to obtain 2000 MCMC samples, and *msm* took 1 hour to obtain 1000 bootstrap samples.

5.3. Reference values and extrapolation

We showed how to derive different types of EC_{10} values using the calibrated joint PBPK-multistate model. We found values in the range of those found with simpler methods. For direct assessment of risk to fish, it might be possible to use external (water) concentration EC_{10s} , taking fish population counts as endpoints. However, for extrapolation of the results to other species, including humans, it would be better to use internal concentration EC_{10s} . The zebrafish PBPK model allowed us to obtain such internal values. Using a human PBPK model (Abduljalil et al., 2019), fetal concentrations could be obtained and the ensuing malformation risk to the embryo would be computed using these internal EC_{10s} . Alternatively, using reverse dosimetry (Brochot et al., 2019), the mother exposure resulting in an acceptable risk level could be computed. A more detailed model could factor in malformations' severity or reversibility and fine-tune reference values and risk assessment.

6. Conclusion

The malformation model we propose is very general. It could be applied to many chemicals and many species. The sub-models linking malformation rates and exposure concentration would simply need to be fine-tuned with species- and chemical-specific data. We showed how to do that using Bayesian inference with *GNU MCSim*, but a maximum likelihood approach and other software like R packages could also be used, at the eventual cost of some loss of flexibility in setting up hierarchical models for variability for example.

For the zebrafish embryo test, our model fits our VPA data well and is an obvious refinement, allowing a better use of data. We could differentiate effects of VPA on various malformation steps to a level that would be inaccessible to simpler dose-response analyses. Our framework could also lead to potential saving of animal lives through optimal design or statistical sample size calculations.

This model could be extended by considering more states and differentiating the severity of malformations. Its coupling to a PBPK would become even more relevant as specific organ concentrations could be links to specific organ malformations. We are also aware of the fact that our approach is relatively complex, and needs to be further verified with many other chemicals before it could be used routinely. It should also be possible to simplify it, by using water concentration as a measure of exposure for example. That could be justified assuming proportionality between water and body concentration. In any case, it should serve the scientific and the risk assessment communities in improving data analyses.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 681002 (*Eu-ToxRisk*). The authors would like to thank Dr. Rachel Rose for useful comments and discussion.

References

- Abduljalil, K., Pan, X., Pansari, A., Jamei, M., Johnson, T.N., 2019. A preterm physiologically based pharmacokinetic model. Part I: physiological parameters and model building. *Clinical Pharmacokinetics*. <https://doi.org/10.1007/s40262-019-00825-6>
- Andersen, P.K., Geskus, R.B., de Witte, T., Putter, H., 2012. Competing risks in epidemiology: possibilities and pitfalls. *International Journal of Epidemiology* 41, 861–870. <https://doi.org/10.1093/ije/dyr213>
- Battistoni, M., Di Renzo, F., Menegola, E., Bois, F.Y., 2019. Quantitative AOP based teratogenicity prediction for mixtures of azole fungicides. *Computational Toxicology* 11, 72–81. <https://doi.org/10.1016/j.comtox.2019.03.004>
- Bernillon, P., Bois, F.Y., 2000. Statistical issues in toxicokinetic modeling: a Bayesian perspective. *Environmental Health Perspectives* 108 (suppl. 5), 883–893.
- Beyersmann, J., Allignol, A., Schumacher, M., 2012. *Competing Risks and Multistate Models with R*. Springer New York, New York, NY. <https://doi.org/10.1007/978-1-4614-2035-4>
- Bois, F.Y., 2012. Bayesian inference, in: Reisfeld, B., Mayeno, A.N. (Eds.), *Computational Toxicology Vol. II*. Humana Press, New-York, pp. 597–636.
- Bois, F.Y., 2009. GNU MCSim: Bayesian statistical inference for SBML-coded systems biology models. *Bioinformatics* 25, 1453–1454. <https://doi.org/10.1093/bioinformatics/btp162>
- Bois, F.Y., Smith, T.J., Gelman, A., Chang, H.Y., Smith, A.E., 1999. Optimal design for a study of butadiene toxicokinetics in humans. *Toxicological Sciences: An Official Journal of the Society of Toxicology* 49, 213–224.
- Bradburn, M.J., Clark, T.G., Love, S.B., Altman, D.G., 2003a. Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods. *British Journal of Cancer* 89, 431–436. <https://doi.org/10.1038/sj.bjc.6601119>
- Bradburn, M.J., Clark, T.G., Love, S.B., Altman, D.G., 2003b. Survival Analysis Part III: Multivariate data analysis – choosing a model and assessing its adequacy and fit. *British Journal of Cancer* 89, 605–611. <https://doi.org/10.1038/sj.bjc.6601120>
- Braunbeck, T., Boettcher, M., Hollert, H., Kosmehl, T., Lammer, E., Leist, E., Rudolf, M., Seitz, N., 2005. Towards an alternative for the acute fish LC(50) test in chemical assessment: the fish embryo toxicity test goes multi-species - an update. *ALTEX* 22, 87–102.
- Braunbeck, T., Kais, B., Lammer, E., Otte, J., Schneider, K., Stengel, D., Strecker, R., 2015. The fish embryo test (FET): origin, applications, and future. *Environmental Science and Pollution Research* 22, 16247–16261. <https://doi.org/10.1007/s11356-014-3814-7>
- Braunbeck, T., Lammer, E., 2006. Draft detailed review paper on fish embryo toxicity assays. UBA Contract Number 203 85 422. Umweltbundesamt – German Federal Environment Agency, Dessau, Germany.
- Brinks, R., Hoyer, A., 2018. Illness-death model: statistical perspective and differential equations. *Lifetime Data Anal* 24, 743–754. <https://doi.org/10.1007/s10985-018-9419-6>
- Brochot, C., Casas, M., Manzano-Salgado, C., Zeman, F.A., Schettgen, T., Vrijheid, M., Bois, F.Y., 2019. Prediction of maternal and foetal exposures to perfluoroalkyl compounds in a

- Spanish birth cohort using toxicokinetic modelling. *Toxicology and Applied Pharmacology* 379, 114640. <https://doi.org/10.1016/j.taap.2019.114640>
- Broetzmann, K., Kroese, D., Wolterbeek, A., Braunbeck, T., submitted. Neurotoxic effects in zebrafish embryos by valproic acid and 9 of its analogues – the fish-mouse connection?
- Busquet, F., Strecker, R., Rawlings, J.M., Belanger, S.E., Braunbeck, T., Carr, G.J., Cenijn, P., Fochtman, P., Gourmelon, A., Hübner, N., Kleensang, A., Knöbel, M., Kussatz, C., Legler, J., Lillicrap, A., Martínez-Jerónimo, F., Polleichtner, C., Rodeczko, H., Salinas, E., Schneider, K.E., Scholz, S., van den Brandhof, E.-J., van der Ven, L.T.M., Walter-Rohde, S., Weigt, S., Witters, H., Halder, M., 2014. OECD validation study to assess intra- and inter-laboratory reproducibility of the zebrafish embryo toxicity test for acute aquatic toxicity testing. *Regulatory Toxicology and Pharmacology* 69, 496–511. <https://doi.org/10.1016/j.yrtph.2014.05.018>
- Cannon, J., Roberts, K., Milne, C., Carapetis, J.R., 2017. Rheumatic heart disease severity, progression and outcomes: a multi-state model. *Journal of the American Heart Association* 6. <https://doi.org/10.1161/JAHA.116.003498>
- Chuang, C.-M., Chang, C.-H., Wang, H.-E., Chen, K.-C., Peng, C.-C., Hsieh, C.-L., Peng, R.Y., 2012. Valproic Acid Downregulates RBP4 and Elicits Hypervitaminosis A-Teratogenesis—A Kinetic Analysis on Retinol/Retinoic Acid Homeostatic System. *PLoS ONE* 7, e43692. <https://doi.org/10.1371/journal.pone.0043692>
- Clark, T.G., Bradburn, M.J., Love, S.B., Altman, D.G., 2003a. Survival Analysis Part I: Basic concepts and first analyses. *British Journal of Cancer* 89, 232–238. <https://doi.org/10.1038/sj.bjc.6601118>
- Clark, T.G., Bradburn, M.J., Love, S.B., Altman, D.G., 2003b. Survival Analysis Part IV: Further concepts and methods in survival analysis. *British Journal of Cancer* 89, 781–786. <https://doi.org/10.1038/sj.bjc.6601117>
- Embry, M.R., Belanger, S.E., Braunbeck, T.A., 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. *Aquatic Toxicology* 97, 79–87. <https://doi.org/10.1016/j.aquatox.2009.12.008>
- EU, 2010. Directive 2010/63/EU of the European parliament and of the council of 22 September 2010 on the protection of animals used for scientific purposes. *Off. J. EU L* 276, 33–79.
- Farewell, V.T., Tom, B.D.M., 2014. The versatility of multi-state models for the analysis of longitudinal data with unobservable features. *Lifetime Data Analysis* 20, 51–75. <https://doi.org/10.1007/s10985-012-9236-2>
- Fathe, K., Palacios, A., Finnell, R.H., 2014. Brief report novel mechanism for valproate-induced teratogenicity: Novel Mechanism for Valproate-Induced Teratogenicity. *Birth Defects Research Part A: Clinical and Molecular Teratology* 100, 592–597. <https://doi.org/10.1002/bdra.23277>
- Fisz, M., 1976. *Probability Theory and Mathematical Statistics*, 3rd ed. Wiley, New York.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science* 7, 457–511.
- Gould, A.L., Boye, M.E., Crowther, M.J., Ibrahim, J.G., Quartey, G., Micallef, S., Bois, F.Y., 2015. Responses to discussants of ‘Joint modeling of survival and longitudinal non-survival

- data: current methods and issues. report of the DIA Bayesian joint modeling working group.' *Statistics in Medicine* 34, 2202–2203. <https://doi.org/10.1002/sim.6502>
- Hill, A.J., Teraoka, H., Heideman, W., Peterson, R.E., 2005. Zebrafish as a Model Vertebrate for Investigating Chemical Toxicity. *Toxicological Sciences* 86, 6–19. <https://doi.org/10.1093/toxsci/kfi110>
- 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. *Journal of Soils and Sediments* 3, 197–207. <https://doi.org/10.1065/jss2003.09.085>
- ISO, 1996. International Organization for Standardization. Water quality - Determination of the 28 acute lethal toxicity of substances to a freshwater fish [*Brachydanio rerio* Hamilton-Buchanan 29 (Teleostei, Cyprinidae)]. ISO 7346-3: Flow-through method. Available: [<http://www.iso.org>].
- Jackson, C.H., 2011. Multi-State Models for Panel Data: The msm Package for R. *J. Stat. Soft.* 38. <https://doi.org/10.18637/jss.v038.i08>
- Jiang, H., Fine, J.P., 2007. Survival Analysis, in: Ambrosius, W.T. (Ed.), *Topics in Biostatistics*. Humana Press, Totowa, NJ, pp. 303–318. https://doi.org/10.1007/978-1-59745-530-5_15
- Jones, E., Epstein, D., García-Mochón, L., 2017. A procedure for deriving formulas to convert transition rates to probabilities for multistate markov models. *Medical Decision Making* 37, 779–789. <https://doi.org/10.1177/0272989X17696997>
- Kantae, V., Krekels, E.H.J., Ordas, A., González, O., van Wijk, R.C., Harms, A.C., Racz, P.I., van der Graaf, P.H., Spaink, H.P., Hankemeier, T., 2016. Pharmacokinetic modeling of paracetamol uptake and clearance in zebrafish larvae: expanding the allometric scale in vertebrates with five orders of magnitude. *Zebrafish* 13, 504–510. <https://doi.org/10.1089/zeb.2016.1313>
- Karmen, C., Gietzelt, M., Knaup-Gregori, P., Ganzinger, M., 2019. Methods for a similarity measure for clinical attributes based on survival data analysis. *BMC Medical Informatics and Decision Making* 19, 195. <https://doi.org/10.1186/s12911-019-0917-6>
- Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., Schilling, T.F., 1995. Stages of embryonic development of the zebrafish. *Developmental Dynamics* 203, 253–310. <https://doi.org/10.1002/aja.1002030302>
- Lammer, E., Carr, G.J., Wendler, K., Rawlings, J.M., Belanger, S.E., Braunbeck, Th., 2009. Is the fish embryo toxicity test (FET) with the zebrafish (*Danio rerio*) a potential alternative for the fish acute toxicity test? *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 149, 196–209. <https://doi.org/10.1016/j.cbpc.2008.11.006>
- Larisch, W., Brown, T.N., Goss, K.-U., 2017. A toxicokinetic model for fish including multiphase sorption features: A high-detailed, physiologically based toxicokinetic model. *Environmental Toxicology and Chemistry* 36, 1538–1546. <https://doi.org/10.1002/etc.3677>
- Lee, E.T., Go, O.T., 1997. Survival analysis in public health research. *Annual Review of Public Health* 18, 105–134. <https://doi.org/10.1146/annurev.publhealth.18.1.105>
- MacRae, C.A., Peterson, R.T., 2015. Zebrafish as tools for drug discovery. *Nature Reviews Drug Discovery* 14, 721–731. <https://doi.org/10.1038/nrd4627>

- Meira-Machado, L., de Uña-Álvarez, J., Cadarso-Suárez, C., Andersen, P.K., 2009. Multi-state models for the analysis of time-to-event data. *Statistical Methods in Medical Research* 18, 195–222. <https://doi.org/10.1177/0962280208092301>
- Nagel, R., 2002. DarT: the embryo testa with the zebrafish *Danio rerio* - A general model in ecotoxicology and toxicology. *Alternatives to Laboratory Animals* 19 Suppl 1, 38–48.
- OECD, 2013. OECD Guidelines for the Testing of Chemicals. Section 2: Effects on Biotic Systems Test No. 236: Fish Embryo Acute Toxicity (FET) Test. Organization for Economic Cooperation and Development, Paris, France. <https://doi.org/10.1787/9789264203709-en>
- Overstall, A.M., Woods, D.C., Parker, B.M., 2020. Bayesian optimal design for ordinary differential equation models with application in biological science. *Journal of the American Statistical Association* 115, 583–598. <https://doi.org/10.1080/01621459.2019.1617154>
- 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. *Journal of Biological Chemistry* 276, 36734–36741. <https://doi.org/10.1074/jbc.M101287200>
- Putter, H., Fiocco, M., Geskus, R.B., 2007. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine* 26, 2389–2430. <https://doi.org/10.1002/sim.2712>
- Quignot, N., Hamon, J., Bois, F.Y., 2014. Extrapolating *in vitro* results to predict human toxicity, in: Bal-Price, A., Jennings, P. (Eds.), *In Vitro Toxicology Systems, Methods in Pharmacology and Toxicology*. Springer Science, New-York, pp. 531–550.
- R Development Core Team, 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Russell, W.M.S., Burch, R.L., 1959. The principles of humane experimental technique. Methuen, London.
- Siméon, S., Brotzmann, K., Fisher, C., Gardner, I., Silvester, S., Maclellann, R., Walker, P., Braunbeck, T., Bois, F.Y., 2020. Development of a generic zebrafish embryo PBPK model and application to the developmental toxicity assessment of valproic acid analogs. *Reproductive Toxicology* 93, 219–229. <https://doi.org/10.1016/j.reprotox.2020.02.010>
- Sipes, N.S., Padilla, S., Knudsen, T.B., 2011. Zebrafish as an integrative model for twenty-first century toxicity testing. *Birth Defects Research Part C: Embryo Today: Reviews* 93, 256–267. <https://doi.org/10.1002/bdrc.20214>
- Smith, A.F.M., Roberts, G.O., 1993. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society Series B* 55, 3–23.
- 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. *Reproductive Toxicology* 33, 128–132. <https://doi.org/10.1016/j.reprotox.2011.06.121>
- Tsiros, P., Bois, F.Y., Dokoumetzidis, A., Tsiliki, G., Sarimveis, H., 2019. Population pharmacokinetic reanalysis of a Diazepam PBPK model: a comparison of Stan and GNU

MCSim. *Journal of Pharmacokinetics and Pharmacodynamics* 46, 173–192. <https://doi.org/10.1007/s10928-019-09630-x>

van der Vaart, M., Spaink, H.P., Meijer, A.H., 2012. Pathogen Recognition and Activation of the Innate Immune Response in Zebrafish. *Advances in Hematology* 2012, 1–19. <https://doi.org/10.1155/2012/159807>

Welton, N.J., Ades, A.E., 2005. Estimation of Markov chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis, and model calibration. *Medical Decision Making* 25, 633–645. <https://doi.org/10.1177/0272989X05282637>

Williams, J.P., Storlie, C.B., Therneau, T.M., Jr, C.R.J., Hannig, J., 2020. A Bayesian approach to multistate hidden Markov models: application to dementia progression. *Journal of the American Statistical Association* 115, 16–31. <https://doi.org/10.1080/01621459.2019.1594831>

Zare, A., Mahmoodi, M., Mohammad, K., Zeraati, H., Hosseini, M., Naieni, K.H., 2014. Assessing Markov and Time Homogeneity Assumptions in Multi-state Models: Application in Patients with Gastric Cancer Undergoing Surgery in the Iran Cancer Institute. *Asian Pacific Journal of Cancer Prevention* 15, 441–447. <https://doi.org/10.7314/APJCP.2014.15.1.44>

DISCUSSION GENERALE ET PERSPECTIVES

Depuis quelques années, les mentalités scientifiques ne cessent d'évoluer concernant les expérimentations animales pour l'évaluation des risques humains, et ce pour diverses raisons. En effet, il leur est souvent reproché de sacrifier à notre bien-être un nombre important d'animaux, d'être longues, coûteuses et de ne pas suffisamment bien prédire la toxicité humaine. De nouveaux outils *in vitro* et *in silico* sont donc régulièrement proposés pour résoudre le problème des 3R (remplacer, réduire, ou raffiner les expérimentations sur animaux) et conforter les résultats des tests *in vivo*. Le projet européen EU-ToxRisk, qui a financé nos travaux, vise entre autres à améliorer les tests de toxicité sur embryons de poisson zèbre, comparables à des tests *in vitro* au regard de la Loi et par leur configuration. Une étude de cas de EU-ToxRisk s'est focalisée sur la toxicité développementale de l'acide valproïque (VPA) et de ses analogues. Dans le cadre de ce projet, nous avons développé deux modèles mathématiques mécanistes qui devraient améliorer la transposabilité à l'Homme des données obtenues chez l'embryon de poisson zèbre.

Le premier est un modèle PBPK générique basé sur la physiologie de l'embryon et capable de prédire les concentrations de très nombreux composés neutres ou ionisables dans ses organes. Dans ce modèle, les coefficients de partage entre le milieu de culture et les organes ou organelles des produits dépendent de leurs propriétés physico-chimiques et sont prédits par l'intermédiaire d'un sous-modèle QSAR, appelé VIVD, au développement duquel nous avons participé. Par ailleurs, le volume des organes de l'embryon de poisson zèbre change très rapidement au cours des cinq premiers jours de son existence et nous en avons tenu compte par un sous-modèle calibré de façon bayésienne à des mesures du volume total embryonnaire. Simplement sur la base, des propriétés du VPA et de ses analogues, le modèle prédit leur distribution dans l'embryon de poisson zèbre avec une erreur relative médiane de l'ordre d'un facteur 0,34. Ensuite, nous nous sommes interrogés sur les paramètres à calibrer pour améliorer le pouvoir prédictif de notre modèle pour ces mêmes produits. Initialement, l'idée était de calibrer leur clairance métabolique puisque d'après la littérature, l'embryon de poisson zèbre est doté de processus métaboliques (de Souza Anselmo et al., 2018b). Cependant, les coefficients de partage étant également prédits, nous avons pris le parti de leur assigner un facteur de correction afin de corriger les biais éventuels du modèle VIVD. *A posteriori*, après calibration bayésienne, il s'avère que la clairance métabolique des analogues du VPA semble

très faible (ce qu'il est important de savoir), et que le modèle VIVD mérite d'être corrigé pour certains composés, quand des données expérimentales le permettent. Pour vérifier les performances prédictives du modèle, qui nous importent le plus, nous l'avons utilisé pour simuler des données issues de la littérature, sur les quantités totales embryonnaires de 17 composés. Puis, nous avons démontré l'application de notre modèle PBPK, capable de prédire les concentrations de produits dans l'embryon, à la correction des concentrations efficaces (EC_{10}) naïves déterminées sur la base des concentrations nominales. Notre modèle permet de relier les effets aux concentrations internes plutôt que nominales, approche beaucoup plus pertinente pour l'étude et la prédiction des malformations. Finalement nous l'avons couplé au modèle pharmacodynamique (PD) multi-états que nous avons développé par la suite.

Notre deuxième contribution est précisément ce modèle multi-états markovien, capable de prédire l'occurrence de malformations en fonction du temps et de la concentration d'exposition en VPA. Nous avons regroupé les 55 malformations observées en cinq états. Par ajustement bayésien du modèle au nombre de transitions observées, nous avons estimé les différents taux de transition entre états (fonctions du temps et des concentrations internes de produits, par couplage au modèle PBPK précédent) et déduit les probabilités de transition en fonction du temps et des concentrations d'exposition. Finalement, pour montrer l'utilité de notre modèle pour l'évaluation des risques, nous avons déterminé des EC_{10} à partir des taux ou des probabilités de transition, pour les comparer aux valeurs observées expérimentalement.

Dans ce chapitre, nous discuterons des principales étapes de développement des deux modèles et les problèmes rencontrés, de leurs limites, ainsi que des points d'amélioration possibles.

1. L'importance des données expérimentales pour un modèle prédictif pertinent

Avant d'entreprendre une étude de toxicité, la cinétique et la dynamique des composés testés devraient être investiguées grâce aux données de la littérature. Cela permet d'émettre des hypothèses à tester, d'avoir une idée du comportement du composé et de proposer un modèle provisoire. La Figure 15 illustre la place essentielle occupée par les tests expérimentaux dans le processus à suivre pour mener à bien un projet de modélisation (Box, 1980). Nous avons fait l'expérience de ses différentes étapes et avons quelques commentaires à leur sujet.

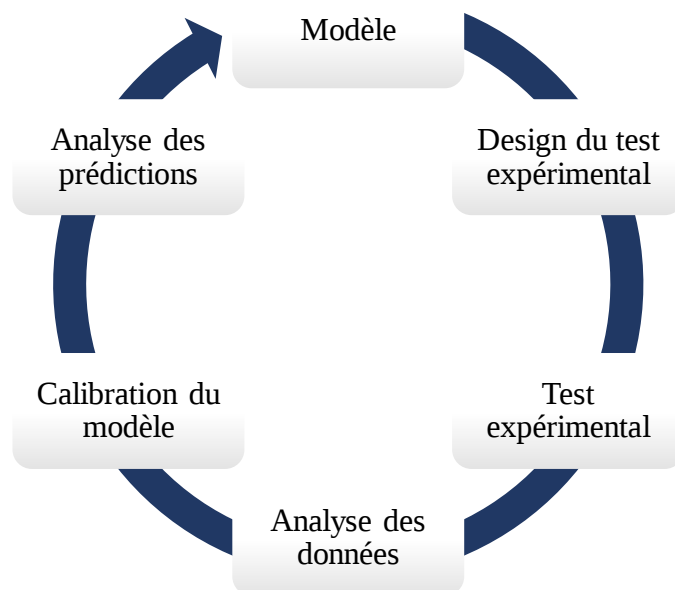


Figure 15. Le carrousel de la modélisation, d'après Gabrielsson et Weiner (2006).

1.1. L'élaboration du test expérimental

La réalisation d'une étude bien conçue est une étape primordiale. Le choix des doses d'exposition, ainsi que les temps de mesures des concentrations internes et des réponses font partie intégrante de l'étape de design du test expérimental. Certaines des données que nous avons utilisées pour l'élaboration du modèle multi-états ont été obtenues selon le protocole de l'OCDE préconisant de relever les malformations d'un nombre minimum requis d'embryons de poisson zèbre à cinq temps fixes différents, et pour au moins cinq doses d'exposition du composé testé. Ce design d'étude offre un jeu de données conséquent pour la construction d'un bon modèle de relation dose-réponse. A l'inverse, pour le modèle PBPK, les données sont insuffisantes puisque les mesures de concentrations embryonnaires et du milieu de culture ont été effectuées seulement à la fin du test (sauf pour le VPA, pour lequel une mesure a été faite aussi à 72 heures). Or, pour comprendre la pharmacocinétique d'un composé il est préférable de disposer de données de concentration relevés à plusieurs temps (Gabrielsson and Weiner, 2012). De plus, pour certains composés étudiés, seulement deux concentrations d'exposition ont été testées, dont des concentrations fortes entraînant quasi instantanément la mort des embryons. Ces données faiblement informatives sont difficilement exploitables pour la pharmacocinétique et entraînent de fortes incertitudes quant au pouvoir prédictif du modèle développé. Nous avons également dû recalibrer notre modèle PBPK suite à une erreur expérimentale. En effet, selon le protocole de l'OCDE en question, le pH de l'eau doit être

vérifié au début du test. Ceci a été fait, mais avant introduction des composés dans l'eau, tamponnée à pH 7.75 avec du chlorure de sodium et de l'acide chlorhydrique. Cependant, un tel tampon n'est pas suffisamment fort pour maintenir le pH constant suite à l'ajout de VPA ou de ses analogues. Après des mesures indépendantes, une diminution du pH du milieu de culture s'est avérée corrélée à l'augmentation des concentrations d'exposition des composés. Or, le pH a un effet important sur l'ionisation des composés étudiés en solution aqueuse, et donc sur leur accumulation et distribution dans l'embryon. Nous avons donc dû recalibrer le modèle PBPK avec les bons pH et écarter l'hypothèse initiale de linéarité de la pharmacocinétique en fonction de l'exposition, puisque les coefficients de partage ne varient pas linéairement avec les concentrations d'exposition. On voit bien l'importance de conduire des tests expérimentaux rigoureux pour l'élaboration d'un modèle performant, car de nombreux facteurs peuvent influencer les suppositions initiales et les prédictions.

1.2. L'analyse des données et ajustement du modèle

Une fois les données recueillies, une analyse exploratoire graphique de données, collectées à partir d'un essai correctement élaboré, est préconisée pour confirmer ou suggérer des modifications du modèle initial. L'étape suivante est la sélection de la forme exacte du modèle et l'ajustement du modèle aux données expérimentales pour estimer ses paramètres et leur précision. Les deux modèles développés ici ont été calibrés par inférence bayésienne. Nos modèles étant nouveaux, nous avons peu d'information *a priori* sur leurs paramètres, et avons donc choisi des priors uniformes. Nos recherches bibliographiques préalables à la calibration du modèle nous ont tout de même donné des informations sur sa structure et nous ont suggéré des valeurs initiales pour les paramètres. Par exemple, pour le modèle PBPK, nous avons imposé des bornes initiales restreintes à la distribution *a priori* de la clairance métabolique, *i.e.* $\sim 0 - 1 \times 10^{-13}$ L/min, puisque des publications démontraient certes l'existence de processus métaboliques chez l'embryon de poisson zèbre mais de façon limitée (Kühnert et al., 2017). Pour le modèle multi-états, les sous-modèles utilisés pour estimer les taux de transitions ont été déduits d'analyses graphiques, ainsi que les valeurs des bornes des distributions *a priori* des paramètres. Cependant, les bornes initialement choisies pour certains paramètres se sont révélées être parfois trop étroites, et nous les avons alors relaxées, ce qui en théorie n'est pas tout à fait correct. Des bornes très larges conduisent à des problèmes de convergence lente des chaînes MCMC partant de régions très éloignées du centre de la distribution *a posteriori*. La sélection du modèle découle généralement d'études antérieures et d'analyses graphiques des

résultats expérimentaux. Par exemple, nous avons testé une hypothèse de non-linéarité du modèle PBPK en essayant une cinétique de Michaelis-Menten pour évaluer la saturabilité du métabolisme. Les ajustements aux données n'étant pas améliorés par rapport à ceux obtenus avec une hypothèse de linéarité, nous avons gardé cette dernière, plus simple, par souci de parcimonie.

1.3. La validation du modèle

Pour établir la capacité du modèle à expliquer et prédire les données, les sorties du modèle, telles que les concentrations embryonnaires pour le modèle PBPK, ou les taux de transition pour le modèle multi-états, doivent être évaluées. Cette étape consiste notamment à déterminer la qualité de l'ajustement, la corrélation entre les paramètres, l'analyse des erreurs résiduelles, ainsi que l'exactitude et la précision des paramètres. Cette évaluation permet également de comparer des modèles entre eux. Les graphiques de diagnostic préconisés représentent les données en fonction des prédictions correspondantes et permettent de statuer visuellement sur le pouvoir prédictif du modèle. Il est également d'usage de douter de la fiabilité du modèle à prédire les données et de le tester sur un nouveau jeu de données. Ce processus est souvent appelé étape de validation croisée du modèle. C'est précisément ce que nous avons fait pour le modèle PBPK avec les données de Brox *et al.* (2016). Cette étape de validation manque au modèle multi-états et mériterait d'être effectuée avant d'entreprendre une complexification de sa structure.

Un modèle peut être *in fine* utilisé soit pour le design de nouvelles expériences, soit plus pragmatiquement pour estimer des paramètres « physiques » fondamentaux, soit pour faire des prédictions utiles. Le but de la construction de modèles est de réduire la quantité de données nécessaires pour l'évaluation de risque, de tester des hypothèses et de comparer des modèles concurrents car un modèle complexe n'est pas forcément signe de performance.

2. Les limites et perspectives d'amélioration des modèles pour prédire la toxicité développementale de l'embryon de poisson zèbre

Les modèles ne représentant pas de façon exhaustive la complexité du système d'intérêt, plusieurs limites leur sont attribuables. Dans une juste mesure, si elles sont pertinemment argumentées et considérées, elles ne conduisent pas obligatoirement à l'invalidité du modèle.

2.1. Le modèle PBPK

Une large incertitude entache des concentrations embryonnaires et du milieu de culture prédites par le modèle PBPK. Les causes en sont multiples et se cumulent : faible précision de la méthode de mesure, variabilité inter-embryonnaire, concentration initiale du milieu de culture différente de la concentration nominale, volatilité possible des composés, dégradation autre que par processus métaboliques, *etc.* Nous allons passer en revue les principales limites connues de ce modèle.

2.1.1. Les coefficients de partage

Les coefficients de partage obtenus par le modèle VIVD constituent les paramètres d'entrée du modèle PBPK. Leurs estimations étant basées sur les propriétés physico-chimiques des composés parfois estimées comme la constante d'Henry, les considérer comme imprécises est justifié. C'est pourquoi nous avons jugé pertinent de leur attribuer un facteur de correction à estimer. Les résultats ont montré qu'il n'était pas nécessaire pour tous les composés étudiés. En revanche, le considérer permet d'atténuer le biais induit par les erreurs d'estimation des coefficients de partage. De plus, le modèle VIVD fait l'hypothèse de diffusion instantanée entre le milieu de culture et l'air, menant probablement à une sur-prédiction de la distribution dans l'air des composés faiblement volatiles. Cependant, modéliser la dynamique de distribution dans l'air, au même titre que la clairance métabolique, complexifierait grandement le modèle et ajouterait de nouveaux paramètres à estimer. La majorité des systèmes d'études des embryons, *i.e.* des plaques de puits, étant généralement hermétiquement couverts, cette amélioration ne nous est pas apparue prioritaire par rapport à l'incertitude cumulée des estimations, d'autant plus que le facteur de correction est également appliqué au coefficient de partage entre l'air et le milieu de culture. Il faut également noter que les propriétés physico-chimiques des tissus, exceptées celles du sac vitellin, sont celles de poissons adultes puisque nous n'avons pas trouvé de données embryonnaires. Des mesures spécifiques seraient intéressantes pour s'affranchir de ce biais et profiter des capacités de prédiction par le modèle VIVD même si à ce jour, la petite taille de l'embryon de poisson zèbre rend cela compromis.

2.1.2. L'hypothèse de diffusion instantanée dans l'embryon

Nous avons supposé une diffusion instantanée des composés au sein de l'embryon, négligeant ainsi la présence du chorion, enveloppe acellulaire de l'embryon, ou la possibilité de transport actif au niveau des parois cellulaires. Cette hypothèse semble raisonnable compte tenu de la petite taille de l'embryon. Cependant, aucune donnée ne nous permet de tester sa validité. Les couches intermédiaires et externes du chorion sont percées de canaux poreux. Malgré de nombreuses spéculations, aucune étude ne permet de clarifier s'il représente une barrière protectrice efficace contre l'exposition aux produits chimiques et s'est avéré ne pas être un obstacle à leur absorption (Brox et al., 2016; Henn and Braunbeck, 2011). Cependant, réaliser les tests de toxicité sur des embryons dépourvus de chorion pourrait supprimer une fraction non négligeable de composé si celui-ci s'y accumule. Cela influencerait l'évaluation du bilan de masse et de l'élimination si la concentration chorionique n'était pas mesurée. Effectuer des mesures de concentrations avec et sans chorion permettrait de mieux comprendre ses effets sur la toxicocinétique.

Des recherches sur la composition des ionocytes, cellules de l'épiderme de l'embryon de poisson zèbre, jouant un rôle important dans l'homéostasie ionique grâce au transport actif des ions transépithéliaux et présentant des similitudes fonctionnelles avec les cellules rénales de mammifères, ont permis de révéler des similitudes entre les transports d'efflux des deux espèces (Gordon et al., 2019). Le transport actif peut jouer un rôle dans la modulation de l'exposition des embryons aux composés, mais nous ne disposons pas de données suffisantes pour le vérifier pour le VPA et ses analogues.

2.1.3. Les volumes tissulaires

L'embryon de poisson zèbre ayant un développement rapide, nous avons considéré primordial d'intégrer la croissance des organes pour prétendre à un modèle PBPK embryonnaire pertinent. Ayant à notre disposition des données de volume embryonnaire total au cours des cinq premiers jours de vie et les fractions volumiques des différents organes au delà du cinquième jour, nous avons développé et paramétré un modèle simple de croissance des organes de l'embryon. Ce modèle tient compte du temps de départ de croissance de chaque organe. Nous avons ajusté les fractions volumiques car la contribution musculaire au volume embryonnaire nous paraissait trop élevée. Ce modèle et ses paramètres sont relativement

empiriques et il serait intéressant de disposer de mesures précises des volumes des organes de l'embryon à différents temps.

2.1.4. La clairance métabolique

Le modèle a déterminé une faible capacité métabolique de l'embryon. Il serait intéressant de pouvoir confirmer cette conclusion par mesure de métabolites attendus, mais cela est difficilement réalisable car des quantités faibles de métabolites se formeraient. Il pourrait être utile de concevoir des expériences sur un court laps de temps avec un faible volume d'eau pour éviter la dilution des métabolites et pouvoir mieux les quantifier. Une autre méthode pourrait consister au dosage de métabolites par utilisation de molécules radiomarquées, comme présenté par l'étude de Fol et al. (2017) chez l'embryon de poisson zèbre âgé de 4 jours en comparaison avec l'adulte pour identifier les différences de processus métaboliques entre les stades de développement.

2.1.5. Les concentrations effectives

L'objectif principal de développer un modèle PBPK pour étudier la toxicité développementale de l'embryon de poisson zèbre était d'estimer les concentrations entraînant 10% des effets sur la base des concentrations prédites internes et non à partir des concentrations nominales, comme cela est fait en routine des tests de toxicité (Figure 16).

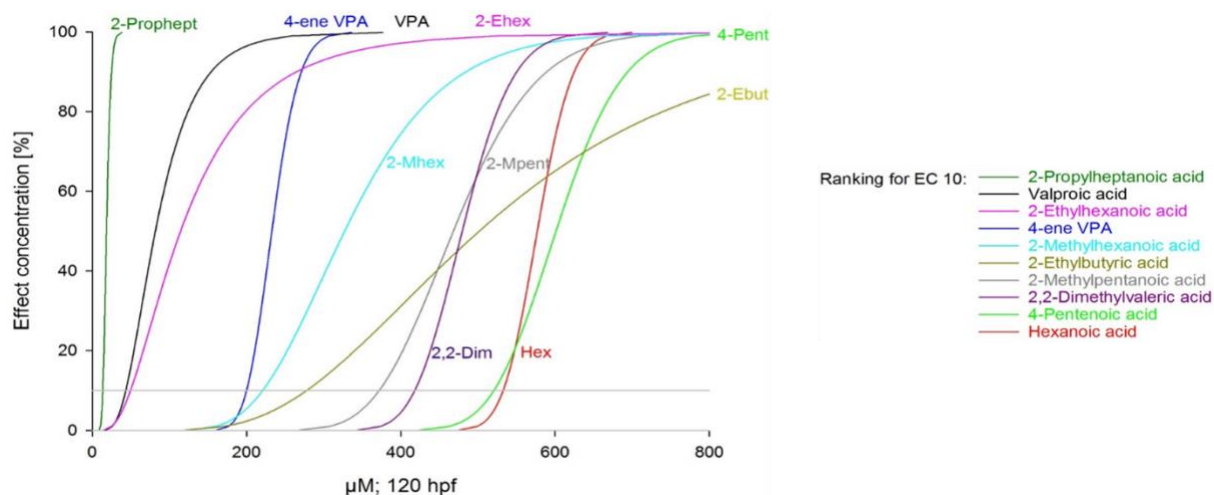


Figure 16. Courbes dose-réponse sur la base des concentrations nominales pour le VPA et ses analogues (d'après les travaux de K. Brotzmann).

Pour l'étude de malformations organiques spécifiques, les concentrations tissulaires prédites devraient être utilisées plutôt que la concentration totale dans l'embryon. Cependant, il faudrait que des mesures de concentrations tissulaires soient disponibles pour valider le modèle. Cela permettrait d'améliorer les interprétations mécanistes pour une meilleure extrapolation inter-espèce et une meilleure évaluation des risques. Les EC_{10} déterminées sur la base des concentrations internes dépendent de la bioaccumulation des composés étudiés. *In fine*, le classement des potentiels toxiques des composés varie et impacte donc les conclusions dans un contexte d'évaluation des risques.

Notre modèle a été appliqué à seulement deux jeux de données : les données du projet EU-ToxRisk et celles publiées par Brox *et al.*(2016) de concentrations à des niveaux d'exposition différents pour un même composé, ce qui est insuffisant pour prétendre à sa validation. Même s'il prédit plutôt bien ces données, nous ne pouvons pas nous fier entièrement à ses prédictions. Davantage de données devraient être collectées pour améliorer le pouvoir prédictif du modèle. Par ces travaux, nous avons démontré que l'ajustement des données peut améliorer les prédictions et qu'en ce sens, un modèle PBPK peut inciter à développer plus de données pour orienter les questions de recherche.

2.1.6. Comparaison avec les modèles de la littérature

Peu de modèles pharmacocinétiques ont été développés pour étudier la toxicité des composés chez l'embryon de poisson zèbre. Il semblerait que nous soyons les premiers à proposer un modèle physiologique. Brox *et al.* (2016a), Nawaji *et al.* (2018b) et Liu *et al.* (2018) utilisent des modèles compartimentaux pour étudier la cinétique et notamment le métabolisme de composés aux propriétés physico-chimiques et caractéristiques toxiques diverses. Pour relier la toxicité développementale à la concentration interne et non à la concentration nominale, Nawaji *et al.* (2018b) étudient avec un modèle pharmacocinétique mono-compartimental l'impact du caractère hydrophobe de trois composés tératogènes, la caféine, le VPA et le diéthylstilbestrol, sur l'évolution de la concentration embryonnaire interne en fonction du temps. Les travaux de Liu *et al.* (2018) concernent l'évaluation de l'accumulation (modèle bi-compartimental) et du métabolisme (modèle mono-compartimental) du tétrabromobisphénol A au cours du développement embryonnaire et larvaire du poisson zèbre. Il s'agit d'un retardateur de flamme dont la structure chimique est similaire à la forme active des hormones thyroïdiennes, soulevant donc des inquiétudes quant à son potentiel toxique sur la fonction thyroïdienne des animaux et des Hommes. Les travaux de Brox *et al.* (2016a) représentent notre

référence principale d'étude pour les améliorations proposées par notre modèle PBPK puisque nos hypothèses de recherches se rejoignent. Nous avons également exploité leurs données pour évaluer la capacité prédictive de notre modèle. Avec un modèle pharmacocinétique mono-compartimental de l'embryon de poisson zèbre, Brox et al. (2016a) ont étudiés l'influence des propriétés physico-chimiques de composés polaires sur l'absorption et l'impact de processus biologiques tels que la biotransformation sur la concentration interne, et donc indirectement sur les effets biologiques résultants. Les ajustements du modèle de la concentration interne ont été effectués selon l'équation suivante :

$$\frac{C_{int}}{C_{ext}} = \frac{k_1}{k_2} (1 - e^{-k_2 t}) \quad (8)$$

Avec C_{int} la concentration interne, C_{ext} la concentration nominale, k_1 le taux d'absorption et k_2 le taux d'élimination. Le rapport k_1/k_2 constitue le facteur cinétique de bioconcentration (BCF). Le modèle considère également l'évolution du volume embryonnaire total pour calculer la concentration interne, estimé à partir de photos microscopiques.

Nous proposons un modèle pharmacocinétique physiologique plutôt qu'un modèle compartimental pour prétendre à une extrapolation inter-espèce pertinente grâce au caractère mécaniste de notre approche. De plus, notre modèle intègre les propriétés physico-chimiques des composés pour définir les coefficients de partage, paramètres d'entrée du modèle PBPK. *A contrario*, pour fournir des informations sur l'influence des propriétés physico-chimiques sur l'absorption des composés par l'embryon de poisson zèbre, le modèle de Brox et al. (2016a) se contente d'être spéculatif quant à leurs différences selon les composés polaires ou apolaires.

Par ajustement bayésien de notre modèle PBPK aux données de quantités embryonnaires totales relevées à 24, 48, 72 et 48 heures post fertilisation, nous avons effectué les prédictions des profils cinétiques pour les 17 composés de l'étude. Nous avons notamment sélectionnés les travaux de Brox et al. (2016a) puisque nous avons le VPA comme composé testé en commun. La comparaison directe des prédictions est faussée en raison des changements du volume embryonnaire puisqu'avec les données d'EU-ToxRisk, les ajustements du modèle sont réalisés avec des données de concentrations embryonnaires totales et non des quantités. De plus, les tests expérimentaux suivis par nos partenaires EU-ToxRisk utilisent des dosages multiples pour harmoniser les concentrations embryonnaires et non un dosage unique au début de l'expérience comme Brox et al. (2016a). Ainsi, nos simulations avaient pour but premier d'évaluer notre modèle avec un jeu de données supplémentaire et non de comparer les deux travaux.

2.2. Le modèle multi-états

Le modèle multi-états markovien que nous avons développé permet en théorie d'utiliser la totalité des informations sur les malformations relevées dans le protocole 236 de l'OCDE. Cependant, s'agissant d'un travail préliminaire, plusieurs limites sont à considérer.

2.2.1. La structure du modèle

Le modèle principal

Pour simplifier le modèle, nous avons regroupé en cinq états seulement les 55 malformations observées pour des intensités et temps différents. Cette réduction pourrait en partie expliquer l'erreur relative des données à faible nombre de transitions. Elle ignore en effet les différences de concentrations tissulaires (pourtant prédites par le modèle PBPK), ainsi que l'évolution différentielle des sensibilités des organes aux malformations. Par exemple, les malformations cardiaques ne peuvent être observées avant le développement du cœur. Aussi, certains types de malformations présentent des niveaux de sévérité évolutifs pour un même organe. Une subdivision organe-spécifique des états rendrait les mesures de l'exposition interne plus pertinentes. Il serait alors nécessaire, bien sûr, de recalibrer le modèle. Le modèle serait d'autant plus mécaniste et l'extrapolation inter-espèce plus prédictive (Cannon et al., 2017). Cependant, scinder les malformations en états entraîne une explosion combinatoire des états associant plusieurs malformations et les embryons polymalformatifs devraient être dans des états spécifiques, à moins de développer une extension du formalisme multi-états. On peut considérer ceci comme une limite structurelle des modèles multi-états et une perspective intéressante de recherche.

Les sous-modèles de transition

Un des événements observés lors du test sur embryon est leur éclosion. Elle correspond à la rupture du chorion. C'est un événement normal du développement embryonnaire, mais l'altération du temps de son occurrence peut être un effet toxique. Notons que le temps d'éclosion pourrait être une donnée censurée par la durée du test de toxicité si ce temps est significativement allongé par l'exposition au toxique (ce qui ne semble pas être le cas pour le VPA). Nous avons discuté ci-dessus des implications éventuelles de la présence du chorion sur la pharmacocinétique des toxiques. Un modèle davantage simplifié à trois états nous a permis

de déterminer la relation entre le temps et le taux normal d'éclosion. Plusieurs fonctions ont été envisagées et un modèle constant par partie a été sélectionné selon le critère d'information d'Akaike (AIC). Il serait pertinent d'affiner ce modèle avec un jeu de données incluant un plus grand nombre de temps d'observation. Il est important de noter que les temps d'observation n'ont pas besoin d'être espacés uniformément dans le cadre d'une approche multi-états. De plus, en raison de l'inclusion, au temps zéro, dans le test de toxicité d'embryons à différents stades précoces non-contrôlés de développement, *i.e.* de 4 à 16 cellules (Figure 17), le temps d'éclosion « normal » varie aléatoirement d'une expérimentation à l'autre. Nous avons adopté une approche hiérarchique pour modéliser cet effet aléatoire. *A posteriori*, les effets du VPA n'apparaissent pas significatifs sur le temps et le taux d'éclosion, et un seul ensemble de paramètres d'éclosion a pu suffire.

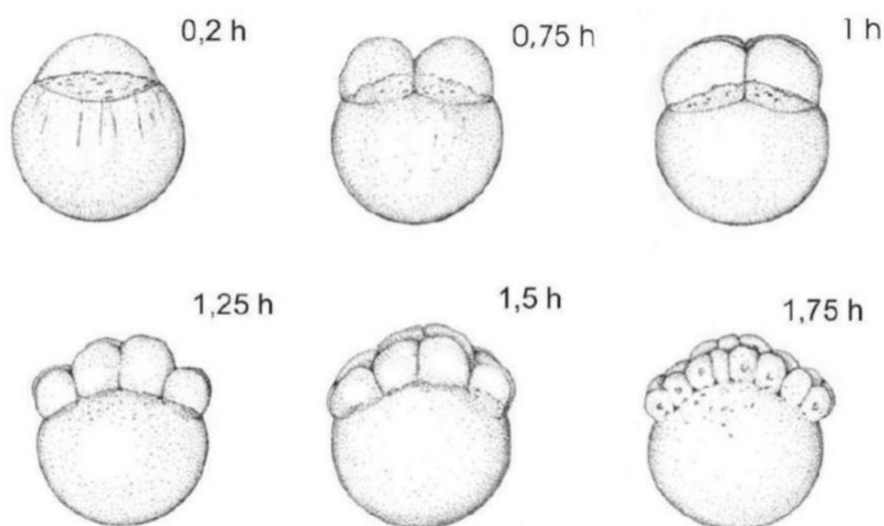


Figure 17. Stades précoces de développement d'inclusion de l'embryon de poisson zèbre pour les tests de toxicité, de 0,2 –1,75 heures post-fertilisation (Kimmel et al., 1995b).

La forme des sous-modèles individuels reliant la concentration interne aux autres taux de transition a été déterminée empiriquement après estimation de ces taux pour chaque groupe d'exposition, indépendamment. Il serait pertinent d'utiliser une méthode plus rigoureuse pour comparer les prédictions des sous-modèles ou un modèle non-paramétrique capable de les décrire. Cependant, le calcul serait considérable et les niveaux de concentrations d'exposition disponibles sont insuffisants pour améliorer et valider les sous-modèles des taux de transition puisque les quatre plus faibles concentrations de VPA testées sont peu effectives. De plus, la concentration de VPA la plus élevée testée, 800 μ M, entraîne la mort de tous les embryons au premier jour, probablement par effet direct sur le pH du milieu de culture. L'effet du pH aurait

pu être modélisé par ajout d'un paramètre spécifique de cette concentration au sous-modèle du taux de transition de coagulation. Cependant, nous avons préféré nous concentrer sur les concentrations causant des malformations et donc exclure les données relatives à cette concentration extrême d'exposition.

Plusieurs transitions inverses d'états pathologiques vers l'état normal ont été observées, estimées et semble être affectées par l'exposition au VPA. Nous aurions pu les considérer comme événement mal classés en utilisant un modèle de Markov à états cachés (Williams et al., 2020), mais les erreurs d'observations semblent difficilement attribuables à l'exposition au VPA. De plus, certaines malformations, comme l'œdème, sont connues pour être réversibles avec la croissance et le développement du système immunitaire de l'embryon (van der Vaart et al., 2012). Par analyse conjointe des données de temps et de concentrations, notre modèle offre l'avantage de caractériser ce type d'évènement peu courant. Ainsi, si les 55 types de malformations étaient analysés, un faible nombre leur serait attribuable, et seul un tel modèle pourrait les quantifier conjointement et de manière adéquate.

2.2.2. La calibration du modèle

En toxicologie, il est d'usage d'exploiter les données sur les nombre d'embryons dans les différents états étudiés en fonction de l'exposition et indépendamment du temps. Ce sont pourtant les transitions qui sont observées. Ces nombres de transition d'un état à un autre sont exactement les données requises pour développer un modèle multi-états. Le nombre d'individus évolue avec le temps et la concentration, mais est moins informatif que le nombre de transition car les trajectoires individuelles à travers les états dans le temps sont perdues. Augmenter la fréquence d'observation requise par le protocole standard de l'OCDE, *i.e.* une fois par jour, serait encore plus informatif. Le modèle considère qu'une transition observée peut se produire par l'intermédiaire d'un autre état, sans observation directe de celui-ci. S'assurer que la plupart des états sont effectivement observés apporterait des prédictions plus précises. De plus, seules les trois doses les plus élevées sont très informatives sur les effets. Notons enfin qu'une application du modèle pourrait être d'estimer la valeur des informations ajoutées et d'aider à planifier d'autres expériences (Bois et al., 1999; Overstall et al., 2020).

Les études publiées et logiciels de modélisation multi-états que nous avons analysés utilisent la solution par exponentielles de matrices, non adaptée aux taux de transition dépendant du temps alors que les modèles multi-états généraux peuvent les intégrer. La solution

traditionnelle à cette limitation auto-imposée, consiste à approximer les paramètres dépendant du temps par des fonctions constantes par parties (Williams et al., 2020). Cependant, cela constitue une limite importante pour lier de façon précise et facile des modèles multi-états à des modèles pharmacocinétiques continus afin de faire des inférences avec un modèle pharmacocinétique-pharmacodynamique. Le nombre de paramètres à estimer est également augmenté. Enfin, la solution par exponentielles de matrices est lourde à calculer. Une solution plus générale consiste à intégrer numériquement l'ensemble des équations différentielles de Kolmogorov. L'intégration numérique n'impose aucune contrainte sur la linéarité des taux de transition ou leur indépendance temporelle. Les algorithmes d'intégration numériques modernes sont très rapides, précis et capable de résoudre des équations fortement non linéaires. Par exemple, dans le domaine de la modélisation thérapeutique, un système multi-états de Kolmogorov pourrait être couplé à un ensemble d'équations différentielles non linéaires pour une concentration de traitement continuellement temps-dépendante et conditionnant les taux de transition.

2.2.3. Les concentrations effectives

Par nos travaux, nous avons montré que le modèle PBPK-multi-états permettait de définir différents types de valeurs EC_{10} . Pour l'évaluation directe du risque chez les poissons, utiliser des EC_{10} dérivées de la concentration du milieu de culture est possible en considérant le nombre d'individus comme paramètre. Cependant, pour l'extrapolation des résultats à d'autres espèces, il est plus pertinent d'exploiter des EC_{10} dérivées de la concentration interne. Par exemple, un modèle PBPK de l'humain (Abduljalil et al., 2019) prédisant des concentrations fœtales pourrait être utilisé pour calculer le risque de malformations pour l'embryon humain via les EC_{10} internes. Alternativement, par dosimétrie inverse (Brochot et al., 2019), l'exposition de la mère entraînant un niveau de risque acceptable pour l'embryon peut également être déterminée. Notre modèle est plus précis et peut tenir compte de la gravité et de la réversibilité des malformations. Il devrait donc permettre d'affiner les valeurs de référence et d'améliorer l'évaluation des risques.

Notre modèle multi-états est très général et pourrait être appliqué à divers composés et à de nombreuses espèces pour une meilleure analyse des données de malformations et aider à concevoir des tests expérimentaux optimaux satisfaisant notamment le principe des 3R. Mais notre approche est relativement complexe et doit être vérifiée avec de nombreux autres produits chimiques avant de prétendre à une utilisation en routine.

3. La place des modèles mécanistes dans l'évaluation des risques

L'identification des dangers environnementaux et l'évaluation des risques des produits chimiques, phytopharmaceutiques, pharmaceutiques, des biocides, des additifs alimentaires et des déchets ne peuvent se dispenser d'expérimentations animales. Les relations entre la dose administrée d'un produit chimique et ses effets peuvent être extrêmement complexes et variées chez l'Homme. Les animaux, en particulier les mammifères, sont assez similaires aux humains en termes de complexité et d'organisation anatomique. Jusqu'à présent, les expériences sur les animaux ont probablement été le meilleur moyen de prédire, par analogie, la réponse humaine en l'absence de données humaines (N. Quignot et al., 2014). Cependant, elles sont au cœur de préoccupations éthiques et économiques, et sont souvent considérées comme inappropriées ou insuffisantes pour évaluer l'ensemble des composés soumis à réglementation. Ainsi, le développement de méthodes alternatives constitue un réel besoin nécessitant des actions de recherches harmonieuses et innovantes de la part des scientifiques de tout domaine. La compréhension mécaniste des processus toxiques ainsi que la détermination de la pharmacocinétique des composés *in vivo* et *in vitro* représentent le point critique pour prétendre à l'élaboration d'approches alternatives pertinentes, tout comme la disponibilité de données *in vivo* de qualité. Ces nécessités sont en faveur au développement de modèles mécanistes PBPK, désormais reconnus et recommandés par des organismes consultatifs scientifiques nationaux et internationaux (EPA, 2006; IPCS, 2010). Pour ces multiples raisons nous avons développé un modèle PBPK de l'embryon de poisson zèbre plutôt qu'un modèle compartimental comme trouvé dans la littérature (Brox et al., 2016a; Liu et al., 2018; Nawaji et al., 2018b), insuffisants pour les substances à cinétique complexe, ou pour réaliser des extrapolations inter-espèces (Grech et al., 2017b). De plus, par approche bayésienne, nos modèles sont capables d'intégrer toutes les données disponibles, permettant alors de considérer de nombreux facteurs avec un rôle potentiel sur le comportement des composés et d'accroître la compréhension mécanistes des processus toxiques

Il est désormais acté que le remplacement des essais sur les animaux par un seul test expérimental ou non est quasiment impossible. Par conséquent, le développement de nouveaux outils *in vitro* et *in silico* représente une valeur ajoutée judicieuse pour améliorer la prédiction de la toxicité des composés et réduire significativement l'utilisation des animaux. Cependant, pour satisfaire l'évaluation des risques, il est nécessaire d'élaborer des approches formalisées permettant à l'industrie et aux instances réglementaires de fournir des informations sur les

dangers potentiels avec fiabilité et support d'un test animal. Il est à noter que pour être approuvée, une méthode scientifique doit obligatoirement être validée et disposer de lignes directrices internationalement acceptées. Un des problèmes majeurs à la mise en place d'approches alternatives en évaluation du risques résulte du fait que leur utilité est en partie hypothétique et pas exclusivement basée sur des arguments scientifiques, causant entrave à la vision objective de leurs avantages et inconvénients. Ainsi, en supplément d'un process de validation solide et d'un raisonnement scientifique convaincant, l'information et l'éducation continue de la communauté scientifique sont indispensables pour créer une meilleure compréhension des opportunités et de l'application potentielle de nouvelles méthodes alternatives (Scholz et al., 2013b). Pour la réalisation de ces travaux de thèse, nous avons collaborés avec des partenaires EU-ToxRisk enclins à l'utilisation des modèles que nous avons développés pour notamment statuer sur les concentrations effectives et proposer des hypothèses quant au comportement des composés. De façon plus générale, les pays membres de l'OCDE s'efforcent à élargir l'utilisation de méthodes alternatives pour l'évaluation des produits chimiques. L'OCDE a élaboré des lignes directrices et des outils pour l'utilisation de méthodes alternatives telles que les (Q)SAR ou les AOP, dans le cadre des Approches intégrées pour tester et évaluer (IATA) (OECD, 2016, 2004). Il est nécessaire d'étudier leur applicabilité en pratique pour différents niveaux de la prise de décision réglementaire, et de s'appuyer sur des études de cas comme cela est fait dans le cadre du projet EU-ToxRisk. L'objectif est de créer une harmonie pour l'utilisation de nouvelles méthodologies et d'orienter de futures recherches par soulèvement d'interrogations scientifiques.

De toute évidence, il est impossible de développer un modèle qui refléterait toute la complexité du système d'étude. La meilleure solution est probablement d'identifier les cibles potentielles de toxicité ou les mécanismes d'un produit chimique via des modèles QSAR ou *in vitro* par exemple, puis de développer un modèle mécaniste ciblé comme nous avons fait pour le modèle PBPK. Suite aux résultats expérimentaux du ZET réalisé par EU-ToxRisk pour étudier la toxicité de l'acide valproïque et de ses analogues, nous avons pris parti de proposer un modèle à dix compartiments pour représenter les organes affectés par l'exposition aux composés. Il est à noter que plusieurs limites évidentes du modèle embryon de poisson zèbre sont à considérer en comparaison aux modèles mammifères, principalement relatives aux barrières physiques (chorion), aux facteurs cinétiques (ADME) et aux facteurs dynamiques (susceptibilité, réactivité). La méthode de dosage utilisée consiste à placer l'embryon dans un milieu de culture contenant une concentration nominale d'exposition prédéterminée.

Cependant, les branchies ne fonctionnant pas et l'embryon étant non autonome à l'alimentation, la principale voie d'exposition au cours des cinq premiers jours de vie est cutanée, ou par échange entre le sac vitellin et l'embryon. La solubilité et les propriétés physico-chimiques des composés peut également biaiser le dosage. De plus, même si les embryons de poisson zèbre semblent être capables de bio-activer certains composés, nous ne savons à ce jour s'ils peuvent métaboliser tous les tératogènes potentiels. Davantage de connaissances sur les caractéristiques ADME de l'embryon de poisson zèbre en développement ou des mesures pour évaluer la dosimétrie interne réduiraient cette limitation. Pour finir, l'exploration des effets peut être altérée par la petite taille de l'embryon puisque sa manipulation est délicate et peut affecter sa viabilité et le comportement cellulaire (Sipes et al., 2011b).

Malgré les limites attribuables à nos modèles et à l'embryon de poisson zèbre, par nos travaux nous proposons une approche alternative, développée et exposée de façon rigoureuse et pertinente, à considérer comme outil utile pour l'évaluation des risques après processus de validation complémentaires. Pour son projet d'études de cas IATA, l'OCDE a d'ailleurs publié l'étude de cas EU-ToxRisk à laquelle nous avons contribué par le développement de notre modèle PBPK de l'embryon de poisson zèbre (<http://www.oecd.org/chemicalsafety/risk-assessment/iata-integrated-approaches-to-testing-and-assessment.htm#Project>).

CONCLUSION GÉNÉRALE

Optimiser la compréhension et la prédiction des processus de toxicité représente un des enjeux prioritaires en évaluation de risques (éco)-toxicologiques. Depuis quelques années, la volonté de satisfaire au principe des 3R (Réduire, Raffiner, Remplacer) a conduit au développement de nouvelles méthodes alternatives à l'expérimentation animale. Dans ce contexte, les approches *in vitro*, *in vivo* et *in silico* concourent à affiner les tests de toxicité, notamment lorsqu'elles sont conjointement utilisées. A travers ces travaux de thèse, nous avons contribué et soutenu ce désir scientifique en proposant des modèles pharmacocinétiques et pharmacodynamiques innovants pour étudier la toxicité développementale de l'acide valproïque et quelques-uns de ses analogues chez l'embryon de poisson zèbre. Malgré les limites qui leur sont attribuables et la nécessité de validations supplémentaires, nombre de leurs avantages peuvent déjà être exploités pour répondre à différentes questions. Dans un premier temps, en utilisant uniquement des données physico-chimiques et des prédictions QSAR du métabolisme, le modèle PBPK peut être utilisé comme un outil *ab initio* à haut débit pour prédire la distribution de composés dans différents tissus embryonnaires, en fonction du temps et des doses d'exposition. De plus, par mise en évidence de limites relatives au manque de données ou aux hypothèses initiales du modèle, il peut orienter de futures recherches et aider à concevoir de nouveaux protocoles expérimentaux. Il permet également de relier les effets aux concentrations organiques et son utilisation devrait améliorer l'extrapolation des données d'embryons de poisson zèbre à l'Homme. Ne disposer que de données de concentrations plasmatiques pour valider des modèles PBPK animaux est monnaie courante et n'empêche en rien leur utilisation intensive dans l'industrie pharmaceutique, y compris dans des contextes réglementaires. Dans un second temps, le modèle multi-états que nous avons développé permet une meilleure analyse et exploitation des informations offertes par les données sur l'occurrence de malformations. Par conception optimale ou calculs statistiques de la taille d'échantillons, il peut contribuer à la réduction du nombre d'animaux nécessaires pour la réalisation des tests de toxicité. Finalement, l'innovation majeure que nos travaux apportent à l'amélioration de l'évaluation de risques, est le couplage des deux modèles permettant de lier les concentrations tissulaires aux malformations spécifiques de certains organes

REFERENCES

- Alderton, W., Berghmans, S., Butler, P., Chassaing, H., Fleming, A., Golder, Z., Richards, F., Gardner, I., 2010. Accumulation and metabolism of drugs and CYP probe substrates in zebrafish larvae. *Xenobiotica* 40, 547–557. <https://doi.org/10.3109/00498254.2010.493960>
- Alzualde, A., Behl, M., Sipes, N.S., Hsieh, J.-H., Alday, A., Tice, R.R., Paules, R.S., Muriana, A., Quevedo, C., 2018. Toxicity profiling of flame retardants in zebrafish embryos using a battery of assays for developmental toxicity, neurotoxicity, cardiotoxicity and hepatotoxicity toward human relevance. *Neurotoxicology and Teratology* 70, 40–50. <https://doi.org/10.1016/j.ntt.2018.10.002>
- Amadoz, A., Sebastian-Leon, P., Vidal, E., Salavert, F., Dopazo, J., 2015. Using activation status of signaling pathways as mechanism-based biomarkers to predict drug sensitivity. *Scientific Reports* 5, 18494. <https://doi.org/10.1038/srep18494>
- Andersen, P.K., Geskus, R.B., de Witte, T., Putter, H., 2012. Competing risks in epidemiology: possibilities and pitfalls. *International Journal of Epidemiology* 41, 861–870. <https://doi.org/10.1093/ije/dyr213>
- Ankley, G.T., Bennett, R.S., Erickson, R.J., Hoff, D.J., Hornung, M.W., Johnson, R.D., Mount, D.R., Nichols, J.W., Russom, C.L., Schmieder, P.K., Serrano, J.A., Tietge, J.E., Villeneuve, D.L., 2010. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry* 29, 730–741. <https://doi.org/10.1002/etc.34>
- ANSES, 2018. Avis de l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail relatif à une demande d'analyse critique des VTR du valproate de sodium (CAS n°1069-66-5).
- Arunachalam, M., Raja, M., Vijayakumar, C., Malaïammal, P., Mayden, R.L., 2013. Natural History of Zebrafish (*Danio rerio*) in India. *Zebrafish* 10, 1–14. <https://doi.org/10.1089/zeb.2012.0803>
- Ashauer, R., Escher, B.I., 2010. Advantages of toxicokinetic and toxicodynamic modelling in aquatic ecotoxicology and risk assessment. *Journal of Environmental Monitoring* 12, 2056. <https://doi.org/10.1039/c0em00234h>
- Bambino, K., Chu, J., 2017. Zebrafish in Toxicology and Environmental Health, in: *Current Topics in Developmental Biology*. Elsevier, pp. 331–367. <https://doi.org/10.1016/bs.ctdb.2016.10.007>
- Banerjee, S., Carlin, B.P., Gelfand, A.E., 2015. Hierarchical modeling and analysis for spatial data, 2. ed. ed, *Monographs on statistics and applied probability*. CRC Press, Boca Raton, Fla.
- Beaudouin, R., Micallef, S., Brochot, C., 2010. A stochastic whole-body physiologically based pharmacokinetic model to assess the impact of inter-individual variability on tissue dosimetry over the human lifespan. *Regulatory Toxicology and Pharmacology* 57, 103–116. <https://doi.org/10.1016/j.yrtph.2010.01.005>
- Becking, G.C., 1995. Use of mechanistic information in risk assessment for toxic chemicals. *Toxicology Letters* 77, 15–24. [https://doi.org/10.1016/0378-4274\(95\)03266-5](https://doi.org/10.1016/0378-4274(95)03266-5)
- Beekhuijzen, M., 2017. The era of 3Rs implementation in developmental and reproductive toxicity (DART) testing: current overview and future perspectives. *Reproductive Toxicology* 72, 86–96. <https://doi.org/10.1016/j.reprotox.2017.05.006>
- Beilmann, M., 2018. Optimizing drug discovery by Investigative Toxicology: Current and future trends. *ALTEX*. <https://doi.org/10.14573/altex.1808181>

- Bellwon, P., Truissi, G.L., Bois, F.Y., Wilmes, A., Schmidt, T., Savary, C.C., Parmentier, C., Hewitt, P.G., Schmal, O., Josse, R., Richert, L., Guillouzo, A., Mueller, S.O., Jennings, P., Testai, E., Dekant, W., 2015. Kinetics and dynamics of cyclosporine A in three hepatic cell culture systems. *Toxicology in Vitro* 30, 62–78. <https://doi.org/10.1016/j.tiv.2015.07.016>
- Benigni, R., Bossa, C., Netzeva, T., Rodomonte, A., Tsakovska, I., 2007. Mechanistic QSAR of aromatic amines: new models for discriminating between homocyclic mutagens and nonmutagens, and validation of models for carcinogens. *Environmental and Molecular Mutagenesis* 48, 754–771. <https://doi.org/10.1002/em.20355>
- Bennet, D., Viswanath, B., Kim, S., An, J.H., 2017. An ultra-sensitive biophysical risk assessment of light effect on skin cells. *Oncotarget* 8, 47861–47875. <https://doi.org/10.18632/oncotarget.18136>
- Bernut, A., Lutfalla, G., Kremer, L., 2015. Regard à travers le danio pour mieux comprendre les interactions hôte/pathogène. *médecine/sciences* 31, 638–646. <https://doi.org/10.1051/medsci/20153106017>
- Beyersmann, J., Allignol, A., Schumacher, M., 2012. *Competing Risks and Multistate Models with R*. Springer New York, New York, NY. <https://doi.org/10.1007/978-1-4614-2035-4>
- Bois, F.Y., 2009. GNU MCSim: Bayesian statistical inference for SBML-coded systems biology models. *Bioinformatics* 25, 1453–1454. <https://doi.org/10.1093/bioinformatics/btp162>
- Bois, F.Y., 2001. Applications of population approaches in toxicology. *Toxicology Letters* 120, 385–394. [https://doi.org/10.1016/S0378-4274\(01\)00270-3](https://doi.org/10.1016/S0378-4274(01)00270-3)
- Borowy, C.S., Ashurst, J.V., 2020. Physiology, zero and first order kinetics, in: *StatPearls*. StatPearls Publishing, Treasure Island (FL).
- Bosak, M., Słowik, A., Turaj, W., 2019. Why do some women with epilepsy use valproic acid despite current guidelines? A single-center cohort study. *Epilepsy & Behavior* 98, 1–5. <https://doi.org/10.1016/j.yebeh.2019.06.031>
- Box, G.E.P., 1980. Sampling and Bayes' inference in scientific modelling and robustness. *Journal of the Royal Statistical Society. Series A (General)* 143, 383. <https://doi.org/10.2307/2982063>
- Bradburn, M.J., Clark, T.G., Love, S.B., Altman, D.G., 2003a. Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods. *British Journal of Cancer* 89, 431–436. <https://doi.org/10.1038/sj.bjc.6601119>
- Bradburn, M.J., Clark, T.G., Love, S.B., Altman, D.G., 2003b. Survival Analysis Part III: Multivariate data analysis – choosing a model and assessing its adequacy and fit. *British Journal of Cancer* 89, 605–611. <https://doi.org/10.1038/sj.bjc.6601120>
- Brannen, K.C., Panzica-Kelly, J.M., Danberry, T.L., Augustine-Rauch, K.A., 2010. Development of a zebrafish embryo teratogenicity assay and quantitative prediction model. *Birth Defects Research Part B: Developmental and Reproductive Toxicology* 89, 66–77. <https://doi.org/10.1002/bdrb.20223>
- Braunbeck, T., Kais, B., Lammer, E., Otte, J., Schneider, K., Stengel, D., Strecker, R., 2015. The fish embryo test (FET): origin, applications, and future. *Environmental Science and Pollution Research* 22, 16247–16261. <https://doi.org/10.1007/s11356-014-3814-7>
- Brinkmann, M., Schlechtriem, C., Reininghaus, M., Eichbaum, K., Buchinger, S., Reifferscheid, G., Hollert, H., Preuss, T.G., 2016. Cross-Species Extrapolation of Uptake and Disposition of Neutral Organic Chemicals in Fish Using a Multispecies Physiologically-Based Toxicokinetic Model Framework. *Environmental Science & Technology* 50, 1914–1923. <https://doi.org/10.1021/acs.est.5b06158>
- Brion, F., Le Page, Y., Piccini, B., Cardoso, O., Tong, S.-K., Chung, B., Kah, O., 2012. Screening estrogenic activities of chemicals or mixtures in vivo using transgenic (cyp19a1b-GFP) zebrafish embryos. *PLoS One* 7, e36069. <https://doi.org/10.1371/journal.pone.0036069>

- Brochot, C., Willemain, M.-É., Zeman, F., 2014. Chapitre 13. La modélisation toxicopharmacocinétique à fondement physiologique : son rôle en évaluation du risque et en pharmacologie, in: *Modéliser & Simuler – Tome 2*. Editions Matériologiques, p. 453. <https://doi.org/10.3917/edmat.varen.2014.01.0453>
- Brox, S., Seiwert, B., Küster, E., Reemtsma, T., 2016a. Toxicokinetics of polar chemicals in zebrafish embryo (*danio rerio*): influence of physicochemical properties and of biological processes. *Environmental Science and Technology* 50, 10264–10272. <https://doi.org/10.1021/acs.est.6b04325>
- Brox, S., Seiwert, B., Küster, E., Reemtsma, T., 2016b. Toxicokinetics of Polar Chemicals in Zebrafish Embryo (*Danio rerio*): Influence of Physicochemical Properties and of Biological Processes. *Environmental Science & Technology* 50, 10264–10272. <https://doi.org/10.1021/acs.est.6b04325>
- Brugman, S., 2016. The zebrafish as a model to study intestinal inflammation. *Developmental & Comparative Immunology* 64, 82–92. <https://doi.org/10.1016/j.dci.2016.02.020>
- Busquet, F., Strecker, R., Rawlings, J.M., Belanger, S.E., Braunbeck, T., Carr, G.J., Cenijn, P., Fochtman, P., Gourmelon, A., Hübner, N., Kleensang, A., Knöbel, M., Kussatz, C., Legler, J., Lillicrap, A., Martínez-Jerónimo, F., Polleichtner, C., Rzedeczek, H., Salinas, E., Schneider, K.E., Scholz, S., van den Brandhof, E.-J., van der Ven, L.T.M., Walter-Rohde, S., Weigt, S., Witters, H., Halder, M., 2014. OECD validation study to assess intra- and inter-laboratory reproducibility of the zebrafish embryo toxicity test for acute aquatic toxicity testing. *Regulatory Toxicology and Pharmacology* 69, 496–511. <https://doi.org/10.1016/j.yrtph.2014.05.018>
- Casati, S., 2018. Integrated Approaches to Testing and Assessment. *Basic Clin Pharmacol Toxicol* 123, 51–55. <https://doi.org/10.1111/bcpt.13018>
- Cassar, S., Adatto, I., Freeman, J.L., Gamse, J.T., Iturria, I., Lawrence, C., Muriana, A., Peterson, R.T., Van Cruchten, S., Zon, L.I., 2020. Use of Zebrafish in Drug Discovery Toxicology. *Chem. Res. Toxicol.* 33, 95–118. <https://doi.org/10.1021/acs.chemrestox.9b00335>
- Chou, W.-C., Lin, Z., 2019. Bayesian evaluation of a physiologically based pharmacokinetic (PBPK) model for perfluorooctane sulfonate (PFOS) to characterize the interspecies uncertainty between mice, rats, monkeys, and humans: development and performance verification. *Environment International* 129, 408–422. <https://doi.org/10.1016/j.envint.2019.03.058>
- Christensen, E.R., Chen, C.-Y., 1985. A general noninteractive multiple toxicity model including probit, logit, and Weibull transformations. *Biometrics* 41, 711. <https://doi.org/10.2307/2531291>
- Christensen, E.R., Kusk, K.O., Nyholm, N., 2009. Dose–response regressions for algal growth and similar continuous endpoints: calculation of effective concentrations. *Environmental Toxicology and Chemistry* 28, 826. <https://doi.org/10.1897/08-068R.1>
- Chuang, C.-M., Chang, C.-H., Wang, H.-E., Chen, K.-C., Peng, C.-C., Hsieh, C.-L., Peng, R.Y., 2012. Valproic Acid Downregulates RBP4 and Elicits Hypervitaminosis A-Teratogenesis—A Kinetic Analysis on Retinol/Retinoic Acid Homeostatic System. *PLoS ONE* 7, e43692. <https://doi.org/10.1371/journal.pone.0043692>
- Clark, T.G., Bradburn, M.J., Love, S.B., Altman, D.G., 2003a. Survival Analysis Part I: Basic concepts and first analyses. *British Journal of Cancer* 89, 232–238. <https://doi.org/10.1038/sj.bjc.6601118>
- Clark, T.G., Bradburn, M.J., Love, S.B., Altman, D.G., 2003b. Survival Analysis Part IV: Further concepts and methods in survival analysis. *British Journal of Cancer* 89, 781–786. <https://doi.org/10.1038/sj.bjc.6601117>

- Crean, D., Bellwon, P., Aschauer, L., Limonciel, A., Moenks, K., Hewitt, P., Schmidt, T., Herrgen, K., Dekant, W., Lukas, A., Bois, F., Wilmes, A., Jennings, P., Leonard, M.O., 2015. Development of an in vitro renal epithelial disease state model for xenobiotic toxicity testing. *Toxicology in Vitro* 30, 128–137. <https://doi.org/10.1016/j.tiv.2014.11.015>
- Csajka, C., Verotta, D., 2006. Pharmacokinetic–Pharmacodynamic Modelling: History and Perspectives. *J Pharmacokinet Pharmacodyn* 33, 227–279. <https://doi.org/10.1007/s10928-005-9002-0>
- D’Costa, A., Shepherd, I.T., 2009. Zebrafish Development and Genetics: Introducing Undergraduates to Developmental Biology and Genetics in a Large Introductory Laboratory Class. *Zebrafish* 6, 169–177. <https://doi.org/10.1089/zeb.2008.0562>
- de Souza Anselmo, C., Sardela, V.F., de Sousa, V.P., Pereira, H.M.G., 2018a. Zebrafish (*Danio rerio*): A valuable tool for predicting the metabolism of xenobiotics in humans? *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 212, 34–46. <https://doi.org/10.1016/j.cbpc.2018.06.005>
- de Souza Anselmo, C., Sardela, V.F., de Sousa, V.P., Pereira, H.M.G., 2018b. Zebrafish (*Danio rerio*): a valuable tool for predicting the metabolism of xenobiotics in humans? *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 212, 34–46. <https://doi.org/10.1016/j.cbpc.2018.06.005>
- Dixit, R., Riviere, J., Krishnan, K., Andersen, M., 2003. Toxicokinetics and physiologically based toxicokinetics in toxicology and risk assessment. *Journal of Toxicology and Environmental Health, Part B* 6, 1–40. <https://doi.org/10.1080/10937400306479>
- Driessen, M., Kienhuis, A.S., Pennings, J.L.A., Pronk, T.E., van de Brandhof, E.-J., Roodbergen, M., Spaink, H.P., van de Water, B., van der Ven, L.T.M., 2013. Exploring the zebrafish embryo as an alternative model for the evaluation of liver toxicity by histopathology and expression profiling. *Arch Toxicol* 87, 807–823. <https://doi.org/10.1007/s00204-013-1039-z>
- Ducrot, V., Ashauer, R., Bednarska, A.J., Hinarejos, S., Thorbek, P., Weyman, G., 2016. Using toxicokinetic-toxicodynamic modeling as an acute risk assessment refinement approach in vertebrate ecological risk assessment: MODELINK Case Study: TK-TD Modeling in Vertebrates. *Integrated Environmental Assessment and Management* 12, 32–45. <https://doi.org/10.1002/ieam.1641>
- Embry, M.R., Belanger, S.E., Braunbeck, T.A., 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. *Aquatic Toxicology* 97, 79–87. <https://doi.org/10.1016/j.aquatox.2009.12.008>
- Engeszer, R.E., Patterson, L.B., Rao, A.A., Parichy, D.M., 2007. Zebrafish in The Wild: A Review of Natural History And New Notes from The Field. *Zebrafish* 4, 21–40. <https://doi.org/10.1089/zeb.2006.9997>
- EPA, 2006. Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment (Final Report). Washington, DC.
- Erhirhie, E.O., Ihekwereme, C.P., Ilodigwe, E.E., 2018. Advances in acute toxicity testing: strengths, weaknesses and regulatory acceptance. *Interdisciplinary Toxicology* 11, 5–12. <https://doi.org/10.2478/intox-2018-0001>
- EU, 2010. Directive 2010/63/EU of the European parliament and of the council of 22 September 2010 on the protection of animals used for scientific purposes. *Off. J. EU L* 276, 33–79.
- European Chemicals Agency, 2016. How to Use Alternatives to Animal Testing to Fulfil the Information Requirements for Reach Registration. Practical Guide. ECHA, Helsinki.
- Farber, S.A., 2001. Genetic Analysis of Digestive Physiology Using Fluorescent Phospholipid Reporters. *Science* 292, 1385–1388. <https://doi.org/10.1126/science.1060418>

- Farewell, V.T., Tom, B.D.M., 2014. The versatility of multi-state models for the analysis of longitudinal data with unobservable features. *Lifetime Data Anal* 20, 51–75. <https://doi.org/10.1007/s10985-012-9236-2>
- Fathe, K., Palacios, A., Finnell, R.H., 2014. Brief report novel mechanism for valproate-induced teratogenicity: Novel Mechanism for Valproate-Induced Teratogenicity. *Birth Defects Research Part A: Clinical and Molecular Teratology* 100, 592–597. <https://doi.org/10.1002/bdra.23277>
- Feng, J., Gao, Y., Chen, M., Xu, X., Huang, M., Yang, T., Chen, N., Zhu, L., 2018. Predicting cadmium and lead toxicities in zebrafish (*Danio rerio*) larvae by using a toxicokinetic–toxicodynamic model that considers the effects of cations. *Science of The Total Environment* 625, 1584–1595. <https://doi.org/10.1016/j.scitotenv.2018.01.068>
- Fisher, C., Siméon, S., Jamei, M., Gardner, I., Bois, Y.F., 2018. VIVD: Virtual in vitro distribution model for the mechanistic prediction of intracellular concentrations of chemicals in in vitro toxicity assays. *Toxicology in Vitro*. <https://doi.org/10.1016/j.tiv.2018.12.017>
- Fleming, A., Diekmann, H., Goldsmith, P., 2013. Functional Characterisation of the Maturation of the Blood-Brain Barrier in Larval Zebrafish. *PLoS ONE* 8, e77548. <https://doi.org/10.1371/journal.pone.0077548>
- Fol, V.L., Brion, F., Hillenweck, A., Perdu, E., Bruel, S., Aït-Aïssa, S., Cravedi, J.-P., Zalko, D., 2017. Comparison of the in vivo biotransformation of two emerging estrogenic contaminants, BP2 and BPs, in zebrafish embryos and adults. *International Journal of Molecular Sciences* 18, 704. <https://doi.org/10.3390/ijms18040704>
- Fry, R.C., Bangma, J., Szilagyi, J., Rager, J.E., 2019. Developing novel in vitro methods for the risk assessment of developmental and placental toxicants in the environment. *Toxicology and Applied Pharmacology* 378, 114635. <https://doi.org/10.1016/j.taap.2019.114635>
- Gabaldón, T., Koonin, E.V., 2013. Functional and evolutionary implications of gene orthology. *Nat Rev Genet* 14, 360–366. <https://doi.org/10.1038/nrg3456>
- Gaballah, S., Swank, A., Sobus, J.R., Howey, X.M., Schmid, J., Catron, T., McCord, J., Hines, E., Strynar, M., Tal, T., 2020. Evaluation of developmental toxicity, developmental neurotoxicity, and tissue dose in zebrafish exposed to GenX and other PFAS. *Environmental Health Perspectives* 128, 047005. <https://doi.org/10.1289/EHP5843>
- Gabrielsson, J., Weiner, D., 2012. Non-compartmental Analysis, in: Reisfeld, B., Mayeno, A.N. (Eds.), *Computational Toxicology*. Humana Press, Totowa, NJ, pp. 377–389. https://doi.org/10.1007/978-1-62703-050-2_16
- Gabrielsson, J., Weiner, D., 2006. *Pharmacokinetic & pharmacodynamic data analysis: concepts and applications*, 4. ed., rev. and expand. ed. Swedish Acedemy of Pharmaceutical Sciences, Stockholm.
- Gao, Y., Zhang, Y., Feng, J., Zhu, L., 2019. Toxicokinetic–toxicodynamic modeling of cadmium and lead toxicity to larvae and adult zebrafish. *Environmental Pollution* S0269749118354046. <https://doi.org/10.1016/j.envpol.2019.05.003>
- Gelman, A., Hennig, C., 2017. Beyond subjective and objective in statistics. *Journal of the Royal Statistical Society Series A* 180, 967–1033. <https://doi.org/10.1111/rssa.12276>
- Goldsmith, J.R., Jobin, C., 2012. Think Small: Zebrafish as a Model System of Human Pathology. *Journal of Biomedicine and Biotechnology* 2012, 1–12. <https://doi.org/10.1155/2012/817341>
- Goldstone, J.V., McArthur, A.G., Kubota, A., Zanette, J., Parente, T., Jönsson, M.E., Nelson, D.R., Stegeman, J.J., 2010. Identification and developmental expression of the full complement of Cytochrome P450 genes in Zebrafish. *BMC Genomics* 11, 643. <https://doi.org/10.1186/1471-2164-11-643>
- Gonçalves, Í.F.S., Souza, T.M., Vieira, L.R., Marchi, F.C., Nascimento, A.P., Farias, D.F., 2020. Toxicity testing of pesticides in zebrafish—a systematic review on chemicals and

- associated toxicological endpoints. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-020-07902-5>
- Gotlib, D., Ramaswamy, R., Kurlander, J.E., DeRiggi, A., Riba, M., 2017. Valproic acid in women and girls of childbearing age. *Current Psychiatry Reports* 19, 58. <https://doi.org/10.1007/s11920-017-0809-3>
- Gould, A.L., Boye, M.E., Crowther, M.J., Ibrahim, J.G., Quartey, G., Micallef, S., Bois, F.Y., 2015. Responses to discussants of ‘Joint modeling of survival and longitudinal non-survival data: current methods and issues. report of the DIA Bayesian joint modeling working group.’ *Statistics in Medicine* 34, 2202–2203. <https://doi.org/10.1002/sim.6502>
- Grech, A., Brochot, C., Dorne, J.-L., Quignot, N., Bois, F.Y., Beaudouin, R., 2017a. Toxicokinetic models and related tools in environmental risk assessment of chemicals. *Science of The Total Environment* 578, 1–15. <https://doi.org/10.1016/j.scitotenv.2016.10.146>
- Grech, A., Brochot, C., Dorne, J.-L., Quignot, N., Bois, F.Y., Beaudouin, R., 2017b. Toxicokinetic models and related tools in environmental risk assessment of chemicals. *Science of The Total Environment* 578, 1–15. <https://doi.org/10.1016/j.scitotenv.2016.10.146>
- Grech, A., Tebby, C., Brochot, C., Bois, F.Y., Bado-Nilles, A., Dorne, J.-L., Quignot, N., Beaudouin, R., 2019. Generic physiologically-based toxicokinetic modelling for fish: integration of environmental factors and species variability. *Science of The Total Environment* 651, 516–531. <https://doi.org/10.1016/j.scitotenv.2018.09.163>
- Groh, K.J., Carvalho, R.N., Chipman, J.K., Denslow, N.D., Halder, M., Murphy, C.A., Roelofs, D., Rolaki, A., Schirmer, K., Watanabe, K.H., 2015. Development and application of the adverse outcome pathway framework for understanding and predicting chronic toxicity: I. Challenges and research needs in ecotoxicology. *Chemosphere* 120, 764–777. <https://doi.org/10.1016/j.chemosphere.2014.09.068>
- Hamilton, F., 1822. An account of the fishes found in the river Ganges and its branches / by Francis Hamilton, (formerly Buchanan,) ...; With a volume of plates in royal quarto. Printed for A. Constable and company; [etc., etc.], Edinburgh : <https://doi.org/10.5962/bhl.title.6897>
- Hammett, L.P., 1937. The effect of structure upon the reactions of organic compounds. benzene derivatives. *Journal of the American Chemical Society* 59, 96–103. <https://doi.org/10.1021/ja01280a022>
- Hansch, C., Maloney, P.P., Fujita, T., Muir, R.M., 1962. Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. *Nature* 194, 178–180. <https://doi.org/10.1038/194178b0>
- Hill, A.J., Teraoka, H., Heideman, W., Peterson, R.E., 2005. Zebrafish as a Model Vertebrate for Investigating Chemical Toxicity. *Toxicological Sciences* 86, 6–19. <https://doi.org/10.1093/toxsci/kfi110>
- Holtzman, N.G., Iovine, M.K., Liang, J.O., Morris, J., 2016. Learning to Fish with Genetics: A Primer on the Vertebrate Model *Danio rerio*. *Genetics* 203, 1069–1089. <https://doi.org/10.1534/genetics.116.190843>
- Hommen, U., Forbes, V., Grimm, V., Preuss, T.G., Thorbek, P., Ducrot, V., 2016. How to use mechanistic effect models in environmental risk assessment of pesticides: Case studies and recommendations from the SETAC workshop MODELINK: MODELINK workshop summary. *Integr Environ Assess Manag* 12, 21–31. <https://doi.org/10.1002/ieam.1704>
- Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J.C., Koch, R., Rauch, G.-J., White, S., Chow, W., Kilian, B., Quintais, L.T., Guerra-Assunção, J.A., Zhou, Y., Gu, Y., Yen, J., Vogel, J.-H., Eyre, T., Redmond, S., Banerjee, R., Chi, J., Fu, B., Langley, E., Maguire, S.F., Laird, G.K., Lloyd, D., Kenyon, E., Donaldson, S., Sehra, H., Almeida-King, J., Loveland, J., Trevanion, S., Jones, M., Quail, M., Willey, D., Hunt, A., Burton, J., Sims, S., McLay, K., Plumb, B., Davis, J., Clee, C., Oliver,

- K., Clark, R., Riddle, C., Elliott, D., Threadgold, G., Harden, G., Ware, D., Mortimer, B., Kerry, G., Heath, P., Phillimore, B., Tracey, A., Corby, N., Dunn, M., Johnson, C., Wood, J., Clark, S., Pelan, S., Griffiths, G., Smith, M., Glithero, R., Howden, P., Barker, N., Stevens, C., Harley, J., Holt, K., Panagiotidis, G., Lovell, J., Beasley, H., Henderson, C., Gordon, D., Auger, K., Wright, D., Collins, J., Raisen, C., Dyer, L., Leung, K., Robertson, L., Ambridge, K., Leongamornlert, D., McGuire, S., Gilderthorp, R., Griffiths, C., Manthravadi, D., Nichol, S., Barker, G., Whitehead, S., Kay, M., Brown, J., Murnane, C., Gray, E., Humphries, M., Sycamore, N., Barker, D., Saunders, D., Wallis, J., Babbage, A., Hammond, S., Mashreghi-Mohammadi, M., Barr, L., Martin, S., Wray, P., Ellington, A., Matthews, N., Ellwood, M., Woodmansey, R., Clark, G., Cooper, J., Tromans, A., Grafham, D., Skuce, C., Pandian, R., Andrews, R., Harrison, E., Kimberley, A., Garnett, J., Fosker, N., Hall, R., Garner, P., Kelly, D., Bird, C., Palmer, S., Gehring, I., Berger, A., Dooley, C.M., Ersan-Ürün, Z., Eser, C., Geiger, H., Geisler, M., Karotki, L., Kirn, A., Konantz, J., Konantz, M., Oberländer, M., Rudolph-Geiger, S., Teucke, M., Osoegawa, K., Zhu, B., Rapp, A., Widaa, S., Langford, C., Yang, F., Carter, N.P., Harrow, J., Ning, Z., Herrero, J., Searle, S.M.J., Enright, A., Geisler, R., Plasterk, R.H.A., Lee, C., Westerfield, M., de Jong, P.J., Zon, L.I., Postlethwait, J.H., Nüsslein-Volhard, C., Hubbard, T.J.P., Crollius, H.R., Rogers, J., Stemple, D.L., 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496, 498–503. <https://doi.org/10.1038/nature12111>
- IPCS, 2010. Characterization and Application of physiologically Based Pharmacokinetic Models in Risk Assessment (Geneva, Switzerland).
- Jackson, C.H., 2011. Multi-State Models for Panel Data: The **msm** Package for R. *J. Stat. Soft.* 38. <https://doi.org/10.18637/jss.v038.i08>
- Jaroach, K., Jaroach, A., Bojko, B., 2018. Cell cultures in drug discovery and development: The need of reliable in vitro-in vivo extrapolation for pharmacodynamics and pharmacokinetics assessment. *Journal of Pharmaceutical and Biomedical Analysis* 147, 297–312. <https://doi.org/10.1016/j.jpba.2017.07.023>
- Jayasimhan, Praveenraj, Rajendran, K., Sibnarayan, D., Kumar, L., Kumar, S., Jesudoss, R., Ramarao, V., 2017. *Danio rerio* (Actinopterygii: Cypriniformes: Cyprinidae): A new record from Andaman Islands, India. *Acta Ichthyol. Piscat.* 47, 391–396. <https://doi.org/10.3750/AIEP/02242>
- Jiang, H., Fine, J.P., 2007. Survival Analysis, in: Ambrosius, W.T. (Ed.), *Topics in Biostatistics*. Humana Press, Totowa, NJ, pp. 303–318. https://doi.org/10.1007/978-1-59745-530-5_15
- Jones, E., Epstein, D., García-Mochón, L., 2017. A Procedure for Deriving Formulas to Convert Transition Rates to Probabilities for Multistate Markov Models. *Med Decis Making* 37, 779–789. <https://doi.org/10.1177/0272989X17696997>
- Jusko, W.J., 2013. Moving from Basic Toward Systems Pharmacodynamic Models. *Journal of Pharmaceutical Sciences* 102, 2930–2940. <https://doi.org/10.1002/jps.23590>
- Kabir, E.R., Rahman, M.S., Rahman, I., 2015. A review on endocrine disruptors and their possible impacts on human health. *Environmental Toxicology and Pharmacology* 40, 241–258. <https://doi.org/10.1016/j.etap.2015.06.009>
- Kantae, V., Krekels, E.H.J., Ordas, A., González, O., van Wijk, R.C., Harms, A.C., Racz, P.I., van der Graaf, P.H., Spaink, H.P., Hankemeier, T., 2016. Pharmacokinetic Modeling of Paracetamol Uptake and Clearance in Zebrafish Larvae: Expanding the Allometric Scale in Vertebrates with Five Orders of Magnitude. *Zebrafish* 13, 504–510. <https://doi.org/10.1089/zeb.2016.1313>
- Kar, S., Roy, K., Leszczynski, J., 2018. Impact of Pharmaceuticals on the Environment: Risk Assessment Using Qsar Modeling Approach, in: Nicolotti, O. (Ed.), *Computational Toxicology*. Springer New York, New York, NY, pp. 395–443. https://doi.org/10.1007/978-1-4939-7899-1_19

- Karmen, C., Gietzelt, M., Knaup-Gregori, P., Ganzinger, M., 2019. Methods for a similarity measure for clinical attributes based on survival data analysis. BMC Med Inform Decis Mak 19, 195. <https://doi.org/10.1186/s12911-019-0917-6>
- Khazaei, M., Ng, C.A., 2018. Evaluating parameter availability for physiologically based pharmacokinetic (PBPK) modeling of perfluorooctanoic acid (PFOA) in zebrafish. Environmental Science: Processes & Impacts 20, 105–119. <https://doi.org/10.1039/C7EM00474E>
- Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., Schilling, T.F., 1995a. Stages of embryonic development of the zebrafish. Developmental Dynamics 203, 253–310. <https://doi.org/10.1002/aja.1002030302>
- Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., Schilling, T.F., 1995b. Stages of embryonic development of the zebrafish. Developmental Dynamics 203, 253–310. <https://doi.org/10.1002/aja.1002030302>
- Kirk, R.G.W., 2018. Recovering *The Principles of Humane Experimental Technique*: The 3Rs and the Human Essence of Animal Research. Science, Technology, & Human Values 43, 622–648. <https://doi.org/10.1177/0162243917726579>
- Kirla, K.T., Groh, K.J., Poetzsch, M., Banote, R.K., Stadnicka-Michalak, J., Eggen, R.I.L., Schirmer, K., Kraemer, T., 2018. Importance of Toxicokinetics to Assess the Utility of Zebrafish Larvae as Model for Psychoactive Drug Screening Using Meta-Chlorophenylpiperazine (mCPP) as Example. Frontiers in Pharmacology 9. <https://doi.org/10.3389/fphar.2018.00414>
- Kirla, K.T., Groh, K.J., Steuer, A.E., Poetzsch, M., Banote, R.K., Stadnicka-Michalak, J., Eggen, R.I.L., Schirmer, K., Kraemer, T., 2016. Zebrafish larvae are insensitive to stimulation by cocaine: importance of exposure route and toxicokinetics. Toxicological Sciences 154, 183–193. <https://doi.org/10.1093/toxsci/kfw156>
- Klüver, N., Ortmann, J., Paschke, H., Renner, P., Ritter, A.P., Scholz, S., 2014. Transient Overexpression of adh8a Increases Allyl Alcohol Toxicity in Zebrafish Embryos. PLoS ONE 9, e90619. <https://doi.org/10.1371/journal.pone.0090619>
- Knudsen, L.E., Smith, A., Törnqvist, E., Forsby, A., Tähti, H., 2019. Nordic symposium on “toxicology and pharmacology without animal experiments—Will it be possible in the next 10 years?” Basic Clin Pharmacol Toxicol 124, 560–567. <https://doi.org/10.1111/bcpt.13193>
- Krishnan, K., Peyret, T., 2009. Physiologically Based Toxicokinetic (PBTK) Modeling in Ecotoxicology, in: Devillers, J. (Ed.), Ecotoxicology Modeling. Springer US, Boston, MA, pp. 145–175. https://doi.org/10.1007/978-1-4419-0197-2_6
- Kuepfer, L., Niederalt, C., Wendl, T., Schlender, J.-F., Willmann, S., Lippert, J., Block, M., Eissing, T., Teutonico, D., 2016. Applied Concepts in PBPK Modeling: How to Build a PBPK/PD Model. CPT Pharmacometrics Syst Pharmacol 5, 516–531. <https://doi.org/10.1002/psp4.12134>
- Kühnert, A., Vogs, C., Seiwert, B., Aulhorn, S., Altenburger, R., Hollert, H., Küster, E., Busch, W., 2017. Biotransformation in the zebrafish embryo – Temporal gene transcription changes of cytochrome P450 enzymes and internal exposure dynamics of the AhR binding xenobiotic benz[a]anthracene. Environmental Pollution 230, 1–11. <https://doi.org/10.1016/j.envpol.2017.04.083>
- Lanzoni, A., Castoldi, A.F., Kass, G.E., Terron, A., De Seze, G., Bal-Price, A., Bois, F.Y., Delclos, K.B., Doerge, D.R., Fritsche, E., Halldorsson, T., Kolossa-Gehring, M., Hougaard Bennekou, S., Koning, F., Lampen, A., Leist, M., Mantus, E., Rousselle, C., Siegrist, M., Steinberg, P., Tritscher, A., Van de Water, B., Vineis, P., Walker, N., Wallace, H., Whelan, M., Younes, M., 2019. Advancing human health risk assessment. EFSA journal. European Food Safety Authority 17, e170712. <https://doi.org/10.2903/j.efsa.2019.e170712>

- Lapointe, G., Commission de la santé et de la sécurité du travail du Québec, Direction des communications, Commission de la santé et de la sécurité du travail du Québec, Service du répertoire toxicologique, 2004. Notions de Toxicologie. Commission de la santé et de la sécurité du travail, Montréal.
- Larisch, W., Brown, T.N., Goss, K.-U., 2017. A toxicokinetic model for fish including multiphase sorption features: A high-detailed, physiologically based toxicokinetic model. *Environmental Toxicology and Chemistry* 36, 1538–1546. <https://doi.org/10.1002/etc.3677>
- Larisch, W., Goss, K.-U., 2018. Modelling oral up-take of hydrophobic and super-hydrophobic chemicals in fish. *Environmental Science: Processes & Impacts* 20, 98–104. <https://doi.org/10.1039/C7EM00495H>
- Lavielle, M., 2018. Pharmacometrics models with hidden Markovian dynamics. *Journal of Pharmacokinetics and Pharmacodynamics* 45, 91–105. <https://doi.org/10.1007/s10928-017-9541-1>
- Lawrence, C., 2011. Advances in Zebrafish Husbandry and Management, in: *Methods in Cell Biology*. Elsevier, pp. 429–451. <https://doi.org/10.1016/B978-0-12-374814-0.00023-9>
- Lee, E.T., Go, O.T., 1997. SURVIVAL ANALYSIS IN PUBLIC HEALTH RESEARCH. *Annu. Rev. Public Health* 18, 105–134. <https://doi.org/10.1146/annurev.publhealth.18.1.105>
- Leist, M., Ghallab, A., Graepel, R., Marchan, R., Hassan, R., Bennekou, S.H., Limonciel, A., Vinken, M., Schildknecht, S., Waldmann, T., Danen, E., van Ravenzwaay, B., Kamp, H., Gardner, I., Godoy, P., Bois, F.Y., Braeuning, A., Reif, R., Oesch, F., Drasdo, D., Höhme, S., Schwarz, M., Hartung, T., Braunbeck, T., Beltman, J., Vrieling, H., Sanz, F., Forsby, A., Gadaleta, D., Fisher, C., Kelm, J., Fluri, D., Ecker, G., Zdrzil, B., Terron, A., Jennings, P., van der Burg, B., Dooley, S., Meijer, A.H., Willighagen, E., Martens, M., Evelo, C., Mombelli, E., Taboureau, O., Mantovani, A., Hardy, B., Koch, B., Escher, S., van Thriel, C., Cadenas, C., Kroese, D., van de Water, B., Hengstler, J.G., 2017. Adverse outcome pathways: opportunities, limitations and open questions. *Archives of Toxicology* 91, 3477–3505. <https://doi.org/10.1007/s00204-017-2045-3>
- Leung, K.-M., Elashoff, R.M., Afifi, A.A., 1997. CENSORING ISSUES IN SURVIVAL ANALYSIS. *Annu. Rev. Public Health* 18, 83–104. <https://doi.org/10.1146/annurev.publhealth.18.1.83>
- Lieschke, G.J., Currie, P.D., 2007. Animal models of human disease: zebrafish swim into view. *Nat Rev Genet* 8, 353–367. <https://doi.org/10.1038/nrg2091>
- Lipscomb, J.C., Haddad, S., Poet, T., Krishnan, K., 2012. Physiologically-Based Pharmacokinetic (PBPK) Models in Toxicity Testing and Risk Assessment, in: Balls, M., Combes, R.D., Bhogal, N. (Eds.), *New Technologies for Toxicity Testing*. Springer US, New York, NY, pp. 76–95. https://doi.org/10.1007/978-1-4614-3055-1_6
- Liu, H., Ma, Z., Zhang, T., Yu, N., Su, G., Giesy, J.P., Yu, H., 2018. Pharmacokinetics and effects of tetrabromobisphenol a (TBBPA) to early life stages of zebrafish (*Danio rerio*). *Chemosphere* 190, 243–252. <https://doi.org/10.1016/j.chemosphere.2017.09.137>
- Lomba, L., Ribate, M.P., Zuriaga, E., García, C.B., Giner, B., 2019. Acute and subacute effects of drugs in embryos of *Danio rerio*. QSAR grouping and modelling. *Ecotoxicology and Environmental Safety* 172, 232–239. <https://doi.org/10.1016/j.ecoenv.2019.01.081>
- Mackay, D., Fraser, A., 2000. Bioaccumulation of persistent organic chemicals: mechanisms and models. *Environmental Pollution* 110, 375–391. [https://doi.org/10.1016/S0269-7491\(00\)00162-7](https://doi.org/10.1016/S0269-7491(00)00162-7)
- MacRae, C.A., Peterson, R.T., 2015. Zebrafish as tools for drug discovery. *Nat Rev Drug Discov* 14, 721–731. <https://doi.org/10.1038/nrd4627>
- Mahnken, J.D., Chan, W., Freeman, D.H., Freeman, J.L., 2008. Reducing the effects of lead-time bias, length bias and over-detection in evaluating screening mammography: a censored

- bivariate data approach. *Stat Methods Med Res* 17, 643–663. <https://doi.org/10.1177/0962280207087309>
- Margiotta-Casaluci, L., Owen, S.F., Huerta, B., Rodríguez-Mozaz, S., Kugathas, S., Barceló, D., Rand-Weaver, M., Sumpter, J.P., 2016. Internal exposure dynamics drive the Adverse Outcome Pathways of synthetic glucocorticoids in fish. *Sci Rep* 6, 21978. <https://doi.org/10.1038/srep21978>
- 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 antiepileptics drugs in zebrafish embryo and larvae. *Neurotoxicology and Teratology* 66, 17–24. <https://doi.org/10.1016/j.ntt.2018.01.008>
- McCarty, L.S., Borgert, C.J., Posthuma, L., 2018. The regulatory challenge of chemicals in the environment: Toxicity testing, risk assessment, and decision-making models. *Regulatory Toxicology and Pharmacology* 99, 289–295. <https://doi.org/10.1016/j.yrtph.2018.10.001>
- McMullen, P.D., Andersen, M.E., Cholewa, B., Clewell, H.J., Dunnick, K.M., Hartman, J.K., Mansouri, K., Minto, M.S., Nicolas, C.I., Phillips, M.B., Slattery, S., Yoon, M., Clewell, R.A., 2018. Evaluating opportunities for advancing the use of alternative methods in risk assessment through the development of fit-for-purpose in vitro assays. *Toxicology in Vitro* 48, 310–317. <https://doi.org/10.1016/j.tiv.2018.01.027>
- Meira-Machado, L., de Uña-Álvarez, J., Cadarso-Suárez, C., Andersen, P.K., 2009. Multi-state models for the analysis of time-to-event data. *Stat Methods Med Res* 18, 195–222. <https://doi.org/10.1177/0962280208092301>
- Menke, A.L., Spitsbergen, J.M., Wolterbeek, A.P.M., Woutersen, R.A., 2011. Normal Anatomy and Histology of the Adult Zebrafish. *Toxicol Pathol* 39, 759–775. <https://doi.org/10.1177/0192623311409597>
- Mihaljević, I., Popović, M., Žaja, R., Maraković, N., Šinko, G., Smital, T., 2017. Interaction between the zebrafish (*Danio rerio*) organic cation transporter 1 (Oct1) and endo- and xenobiotics. *Aquatic Toxicology* 187, 18–28. <https://doi.org/10.1016/j.aquatox.2017.03.012>
- Moore, D.F., 2016. *Applied Survival Analysis Using R, Use R!* Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-31245-3>
- Mumtaz, M., Fisher, J., Blount, B., Ruiz, P., 2012. Application of physiologically based pharmacokinetic models in chemical risk assessment. *Journal of Toxicology* 2012, 1–11. <https://doi.org/10.1155/2012/904603>
- Munir, A., Azam, S., Fazal, S., Bhatti, A.I., 2018. Evaluation of the whole body physiologically based pharmacokinetic (WB-PBPK) modeling of drugs. *Journal of Theoretical Biology* 451, 1–9. <https://doi.org/10.1016/j.jtbi.2018.04.032>
- Mushak, P., 2013. Limits to chemical hormesis as a dose–response model in health risk assessment. *Science of The Total Environment* 443, 643–649. <https://doi.org/10.1016/j.scitotenv.2012.11.017>
- Nawaji, T., Mizoguchi, N., Ono, M., Matuura, T., Seki, M., Teraoka, H., 2018a. Comparing time-series of chemical concentrations in zebrafish (*Danio rerio*) embryos/larvae exposed to teratogens with different hydrophobicity; caffeine, sodium valproate, and diethylstilbestrol. *The Journal of Toxicological Sciences* 43, 267–273. <https://doi.org/10.2131/jts.43.267>
- Nawaji, T., Mizoguchi, N., Ono, M., Matuura, T., Seki, M., Teraoka, H., 2018b. Comparing time-series of chemical concentrations in zebrafish (*Danio rerio*) embryos/larvae exposed to teratogens with different hydrophobicity; caffeine, sodium valproate, and diethylstilbestrol. *The Journal of Toxicological Sciences* 43, 267–273. <https://doi.org/10.2131/jts.43.267>
- OECD, 2016. *Users’ handbook supplement to the guidance document for developing and assessing AOPs.*
- OECD, 2013. *OECD guidelines for the testing of chemicals. Fish Embryo Acute Toxicity (FET) Test.*

- OECD, 2007. OECD Environment Health and Safety Publications Series on Testing and Assessment No. 80.
- OECD, 2004. OECD principles for the validation, for regulatory purposes, of (quantitative) structure-activity relationship models.
- O'Hagan, A., 2008. The Bayesian Approach to Statistics, in: Handbook of Probability: Theory and Applications. SAGE Publications, Inc., Thousand Oaks, pp. 85–100. <https://doi.org/10.4135/9781452226620.n6>
- Papa, E., van der Wal, L., Arnot, J.A., Gramatica, P., 2014. Metabolic biotransformation half-lives in fish: QSAR modeling and consensus analysis. Science of The Total Environment 470–471, 1040–1046. <https://doi.org/10.1016/j.scitotenv.2013.10.068>
- Parasuraman, S., 2011. Toxicological screening. J Pharmacol Pharmacother 2, 74. <https://doi.org/10.4103/0976-500X.81895>
- Pearce, R.G., Setzer, R.W., Davis, J.L., Wambaugh, J.F., 2017. Evaluation and calibration of high-throughput predictions of chemical distribution to tissues. Journal of Pharmacokinetics and Pharmacodynamics 44, 549–565. <https://doi.org/10.1007/s10928-017-9548-7>
- Peña-Chilet, M., Esteban-Medina, M., Falco, M.M., Rian, K., Hidalgo, M.R., Loucera, C., Dopazo, J., 2019. Using mechanistic models for the clinical interpretation of complex genomic variation. Scientific Reports 9, 18937. <https://doi.org/10.1038/s41598-019-55454-7>
- Perkins, E.J., Ashauer, R., Burgoon, L., Conolly, R., Landesmann, B., Mackay, C., Murphy, C.A., Pollesch, N., Wheeler, J.R., Zupanec, A., Scholz, S., 2019. Building and applying quantitative adverse outcome pathway models for chemical hazard and risk assessment. Environmental Toxicology and Chemistry 38, 1850–1865. <https://doi.org/10.1002/etc.4505>
- Péry, A.R.R., Devillers, J., Brochot, C., Mombelli, E., Palluel, O., Piccini, B., Brion, F., Beaudouin, R., 2014. A Physiologically Based Toxicokinetic Model for the Zebrafish *Danio rerio*. Environmental Science & Technology 48, 781–790. <https://doi.org/10.1021/es404301q>
- Philippe Saint Pierre, 2015. Introduction à l'analyse des durées de survie.
- Popovic, M., Zaja, R., Fent, K., Smital, T., 2014. Interaction of environmental contaminants with zebrafish organic anion transporting polypeptide, Oatp1d1 (Slco1d1). Toxicology and Applied Pharmacology 280, 149–158. <https://doi.org/10.1016/j.taap.2014.07.015>
- Punt, A., 2020. New approach methodologies (NAMs) for human-relevant biokinetics predictions. ALTEX. <https://doi.org/10.14573/altex.2003242>
- Putter, H., Fiocco, M., Geskus, R.B., 2007. Tutorial in biostatistics: competing risks and multi-state models. Statistics in Medicine 26, 2389–2430. <https://doi.org/10.1002/sim.2712>
- Quignot, Nadia, Hamon, J., Bois, F.Y., 2014. Extrapolating in vitro results to predict human toxicity, in: Bal-Price, A., Jennings, P. (Eds.), In Vitro Toxicology Systems. Springer New York, New York, NY, pp. 531–550. https://doi.org/10.1007/978-1-4939-0521-8_24
- Quignot, N., Hamon, J., Bois, F.Y., 2014. Extrapolating *in vitro* results to predict human toxicity, in: Bal-Price, A., Jennings, P. (Eds.), In Vitro Toxicology Systems, Methods in Pharmacology and Toxicology. Springer Science, New-York, pp. 531–550.
- Rahman, A., Munther, D., Fazil, A., Smith, B., Wu, J., 2018. Advancing risk assessment: mechanistic dose–response modelling of *Listeria monocytogenes* infection in human populations. Royal Society Open Science 5, 180343. <https://doi.org/10.1098/rsos.180343>
- Rahmani, S., Breyner, N.M., Su, H.-M., Verdu, E.F., Didar, T.F., 2019. Intestinal organoids: A new paradigm for engineering intestinal epithelium in vitro. Biomaterials 194, 195–214. <https://doi.org/10.1016/j.biomaterials.2018.12.006>
- Raies, A.B., Bajic, V.B., 2016. *In silico* toxicology: computational methods for the prediction of chemical toxicity: computational methods for the prediction of chemical toxicity. Wiley Interdisciplinary Reviews: Computational Molecular Science 6, 147–172. <https://doi.org/10.1002/wcms.1240>

- Reed, M.C., Thomas, R.L., Pavisic, J., James, S.J., Ulrich, C.M., Nijhout, H.F., 2008. A mathematical model of glutathione metabolism. *Theoretical Biology and Medical Modelling* 5, 8. <https://doi.org/10.1186/1742-4682-5-8>
- Rhomberg, L.R., 2010. Toxicity Testing in the 21st Century: How will it Affect Risk Assessment? *Journal of Toxicology and Environmental Health, Part B* 13, 361–375. <https://doi.org/10.1080/10937404.2010.483951>
- Ringeissen, S., Marrot, L., Note, R., Labarussiat, A., Imbert, S., Todorov, M., Mekenyan, O., Meunier, J.-R., 2011. Development of a mechanistic SAR model for the detection of phototoxic chemicals and use in an integrated testing strategy. *Toxicology in Vitro* 25, 324–334. <https://doi.org/10.1016/j.tiv.2010.09.017>
- Russom, C.L., Breton, R.L., Walker, J.D., Bradbury, S.P., 2003. An overview of the use of quantitative structure–activity relationships for ranking and prioritizing large chemical inventories for environmental risk assessments. *Environmental Toxicology and Chemistry* 22, 1810. <https://doi.org/10.1897/01-194>
- Saad, M., Matheeußen, A., Bijttebier, S., Verbueken, E., Pype, C., Casteleyn, C., Van Ginneken, C., Apers, S., Maes, L., Cos, P., Van Cruchten, S., 2017. In vitro CYP-mediated drug metabolism in the zebrafish (embryo) using human reference compounds. *Toxicology in Vitro* 42, 329–336. <https://doi.org/10.1016/j.tiv.2017.05.009>
- Schmidt, U., Struck, S., Gruening, B., Hossbach, J., Jaeger, I.S., Parol, R., Lindequist, U., Teuscher, E., Preissner, R., 2009. SuperToxic: a comprehensive database of toxic compounds. *Nucleic Acids Research* 37, D295–D299. <https://doi.org/10.1093/nar/gkn850>
- Schmieder, P., Mekenyan, O., Bradbury, S., Veith, G., 2003. QSAR prioritization of chemical inventories for endocrine disruptor testing. *Pure and Applied Chemistry* 75, 2389–2396. <https://doi.org/10.1351/pac200375112389>
- Scholz, S., Sela, E., Blaha, L., Braunbeck, T., Galay-Burgos, M., García-Franco, M., Guinea, J., Klüver, N., Schirmer, K., Tanneberger, K., Tobor-Kapłon, M., Witters, H., Belanger, S., Benfenati, E., Creton, S., Cronin, M.T.D., Eggen, R.I.L., Embry, M., Ekman, D., Gourmelon, A., Halder, M., Hardy, B., Hartung, T., Hubesch, B., Jungmann, D., Lampi, M.A., Lee, L., Léonard, M., Küster, E., Lillicrap, A., Luckenbach, T., Murk, A.J., Navas, J.M., Peijnenburg, W., Repetto, G., Salinas, E., Schüürmann, G., Spielmann, H., Tollefsen, K.E., Walter-Rohde, S., Whale, G., Wheeler, J.R., Winter, M.J., 2013a. A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. *Regulatory Toxicology and Pharmacology* 67, 506–530. <https://doi.org/10.1016/j.yrtph.2013.10.003>
- Scholz, S., Sela, E., Blaha, L., Braunbeck, T., Galay-Burgos, M., García-Franco, M., Guinea, J., Klüver, N., Schirmer, K., Tanneberger, K., Tobor-Kapłon, M., Witters, H., Belanger, S., Benfenati, E., Creton, S., Cronin, M.T.D., Eggen, R.I.L., Embry, M., Ekman, D., Gourmelon, A., Halder, M., Hardy, B., Hartung, T., Hubesch, B., Jungmann, D., Lampi, M.A., Lee, L., Léonard, M., Küster, E., Lillicrap, A., Luckenbach, T., Murk, A.J., Navas, J.M., Peijnenburg, W., Repetto, G., Salinas, E., Schüürmann, G., Spielmann, H., Tollefsen, K.E., Walter-Rohde, S., Whale, G., Wheeler, J.R., Winter, M.J., 2013b. A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. *Regulatory Toxicology and Pharmacology* 67, 506–530. <https://doi.org/10.1016/j.yrtph.2013.10.003>
- Schultz, T.W., Cronin, M.T.D., Walker, J.D., Aptula, A.O., 2003. Quantitative structure–activity relationships (QSARs) in toxicology: a historical perspective. *Journal of Molecular Structure: THEOCHEM* 622, 1–22. [https://doi.org/10.1016/S0166-1280\(02\)00614-0](https://doi.org/10.1016/S0166-1280(02)00614-0)
- Sharma, R.P., Schuhmacher, M., Kumar, V., 2017. Developing integrated PBPK/PD coupled mechanistic pathway model (miRNA-BDNF): An approach towards system toxicology. *Toxicology Letters* 280, 79–91. <https://doi.org/10.1016/j.toxlet.2017.08.003>
- Shehwana, H., Konu, O., 2019. Comparative transcriptomics between zebrafish and mammals: a roadmap for discovery of conserved and unique signaling pathways in physiology and

- disease. *Frontiers in Cell and Developmental Biology* 7, 5. <https://doi.org/10.3389/fcell.2019.00005>
- Shin, M.-Y., Shin, C., Choi, J.W., Lee, J., Lee, S., Kim, S., 2019. Pharmacokinetic profile of propyl paraben in humans after oral administration. *Environment International* 130, 104917. <https://doi.org/10.1016/j.envint.2019.104917>
- Siméon, S., Brotzmann, K., Fisher, C., Gardner, I., Silvester, S., Maclellan, R., Walker, P., Braunbeck, T., Bois, F.Y., 2020. Development of a generic zebrafish embryo PBPK model and application to the developmental toxicity assessment of valproic acid analogs. *Reproductive Toxicology* 93, 219–229. <https://doi.org/10.1016/j.reprotox.2020.02.010>
- Sipes, N.S., Padilla, S., Knudsen, T.B., 2011a. Zebrafish-As an integrative model for twenty-first century toxicity testing. *Birth Defects Research Part C: Embryo Today: Reviews* 93, 256–267. <https://doi.org/10.1002/bdrc.20214>
- Sipes, N.S., Padilla, S., Knudsen, T.B., 2011b. Zebrafish as an integrative model for twenty-first century toxicity testing. *Birth Defects Research Part C: Embryo Today: Reviews* 93, 256–267. <https://doi.org/10.1002/bdrc.20214>
- Smith, A.F.M., Roberts, G.O., 1993. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society Series B* 55, 3–23. <https://doi.org/10.1111/j.2517-6161.1993.tb01466.x>
- Sobanska, M., Scholz, S., Nyman, A.-M., Cesnaitis, R., Gutierrez Alonso, S., Klüver, N., Kühne, R., Tyle, H., de Knecht, J., Dang, Z., Lundbergh, I., Carlon, C., De Coen, W., 2018. Applicability of the fish embryo acute toxicity (FET) test (OECD 236) in the regulatory context of Registration, Evaluation, Authorisation, and Restriction of Chemicals (REACH): Fish embryo acute toxicity (FET) test and REACH regulation. *Environmental Toxicology and Chemistry* 37, 657–670. <https://doi.org/10.1002/etc.4055>
- Spence, R., Gerlach, G., Lawrence, C., Smith, C., 2007. The behaviour and ecology of the zebrafish, *Danio rerio*. *Biological Reviews* 83, 13–34. <https://doi.org/10.1111/j.1469-185X.2007.00030.x>
- Spinu, N., Cronin, M.T.D., Enoch, S.J., Madden, J.C., Worth, A.P., 2020. Quantitative adverse outcome pathway (qAOP) models for toxicity prediction. *Archives of Toxicology* 94, 1497–1510. <https://doi.org/10.1007/s00204-020-02774-7>
- Spitsbergen, J.M., Kent, M.L., 2003. The state of the art of the zebrafish model for toxicology and toxicologic pathology research—advantages and current limitations. *Toxicol Pathol* 31 Suppl, 62–87. <https://doi.org/10.1080/01926230390174959>
- Stewart, A.M., Braubach, O., Spitsbergen, J., Gerlai, R., Kalueff, A.V., 2014. Zebrafish models for translational neuroscience research: from tank to bedside. *Trends in Neurosciences* 37, 264–278. <https://doi.org/10.1016/j.tins.2014.02.011>
- 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. *Reproductive Toxicology* 33, 128–132. <https://doi.org/10.1016/j.reprotox.2011.06.121>
- Sztajnkrzycki, M.D., 2002. Valproic acid toxicity: overview and management. *J. Toxicol. Clin. Toxicol.* 40, 789–801.
- Tan, Y.-M., Worley, R.R., Leonard, J.A., Fisher, J.W., 2018. Challenges Associated With Applying Physiologically Based Pharmacokinetic Modeling for Public Health Decision-Making. *Toxicological Sciences* 162, 341–348. <https://doi.org/10.1093/toxsci/kfy010>
- Taylor, K., 2018. Ten Years of REACH — An Animal Protection Perspective. *Alternatives to Laboratory Animals* 46, 347–373. <https://doi.org/10.1177/026119291804600610>
- Urso, R., Blardi, P., Giorgi, G., 2002. A short introduction to pharmacokinetics. *European Review for Medical and Pharmacological Sciences* 6, 33–44.

- van der Vaart, M., Spaink, H.P., Meijer, A.H., 2012. Pathogen Recognition and Activation of the Innate Immune Response in Zebrafish. *Advances in Hematology* 2012, 1–19. <https://doi.org/10.1155/2012/159807>
- van der Worp, H.B., Howells, D.W., Sena, E.S., Porritt, M.J., Rewell, S., O’Collins, V., Macleod, M.R., 2010. Can Animal Models of Disease Reliably Inform Human Studies? *PLoS Med* 7, e1000245. <https://doi.org/10.1371/journal.pmed.1000245>
- Van Norman, G.A., 2019. Limitations of Animal Studies for Predicting Toxicity in Clinical Trials. *JACC: Basic to Translational Science* 4, 845–854. <https://doi.org/10.1016/j.jacbts.2019.10.008>
- Van Wijk, R.C., Krekels, E.H.J., Kantae, V., Ordas, A., Kreling, T., Harms, A.C., Hankemeier, T., Spaink, H.P., van der Graaf, P.H., 2019. Mechanistic and Quantitative Understanding of Pharmacokinetics in Zebrafish Larvae through Nanoscale Blood Sampling and Metabolite Modeling of Paracetamol. *J Pharmacol Exp Ther* 371, 15–24. <https://doi.org/10.1124/jpet.119.260299>
- Varela, M., Dios, S., Novoa, B., Figueras, A., 2012. Characterisation, expression and ontogeny of interleukin-6 and its receptors in zebrafish (*Danio rerio*). *Developmental & Comparative Immunology* 37, 97–106. <https://doi.org/10.1016/j.dci.2011.11.004>
- Vargesson, N., 2015. Thalidomide-induced teratogenesis: history and mechanisms. *Birth Defects Research Part C* 105, 140–156. <https://doi.org/10.1002/bdrc.21096>
- Wakefield, J., Rahman, N., 2000. The combination of population pharmacokinetic studies. *Biometrics* 56, 263–270. <https://doi.org/10.1111/j.0006-341X.2000.00263.x>
- Walker, J., Carlsen, L., Jaworska, J., 2003. Improving opportunities for regulatory acceptance of qsars: the importance of model domain, uncertainty, validity and predictability. *QSAR & Combinatorial Science* 22, 346–350. <https://doi.org/10.1002/qsar.200390024>
- Watson, D.E., Hunziker, R., Wikswo, J.P., 2017. Fitting tissue chips and microphysiological systems into the grand scheme of medicine, biology, pharmacology, and toxicology. *Exp Biol Med (Maywood)* 242, 1559–1572. <https://doi.org/10.1177/1535370217732765>
- WHO, 2010. International Programme on Chemical Safety, Harmonization Project Document No.9. Characterization and Application of Physiologically Based Pharmacokinetic Models in Risk Assessment.
- Williams, J.P., Storlie, C.B., Therneau, T.M., Jr, C.R.J., Hannig, J., 2020. A Bayesian approach to multistate hidden Markov models: application to dementia progression. *Journal of the American Statistical Association* 115, 16–31. <https://doi.org/10.1080/01621459.2019.1594831>
- Zare, A., Mahmoodi, M., Mohammad, K., Zeraati, H., Hosseini, M., Naieni, K.H., 2014. Assessing Markov and Time Homogeneity Assumptions in Multi-state Models: Application in Patients with Gastric Cancer Undergoing Surgery in the Iran Cancer Institute. *Asian Pacific Journal of Cancer Prevention* 15, 441–447. <https://doi.org/10.7314/APJCP.2014.15.1.441>
- Zhao, X., Pack, M., 2017. Modeling intestinal disorders using zebrafish, in: *Methods in Cell Biology*. Elsevier, pp. 241–270. <https://doi.org/10.1016/bs.mcb.2016.11.006>
- Zoupa, M., Zwart, E.P., Gremmer, E.R., Nugraha, A., Compeer, S., Slob, W., van der Ven, L.T.M., 2020. Dose addition in chemical mixtures inducing craniofacial malformations in zebrafish (*Danio rerio*) embryos. *Food and Chemical Toxicology* 137, 111117. <https://doi.org/10.1016/j.fct.2020.111117>

ANNEXES

Annexe 1 : Supplementary Material (Article modèle VIVD)

Formulas for cellular compartment to extracellular medium partition coefficients.

For intracellular water:

$$K_{cell,iw,u} = \frac{C_{cell,iw}}{C_{media,unbound}} = \left(f_{iw} + \frac{P_{nl}f_{nl} + P_{np}f_{np} + K_{aAP}[AP^-](Y_{1,iw} + Y_{2,iw})}{X_{iw}} \right) \frac{X_{iw}}{X_{ew}} K_{cell^{uu,uu}} \quad (27)$$

For lysosomes:

$$K_{lyso,u} = \frac{C_{lyso}}{C_{media,unbound}} = \left(f_{iw} + \frac{P_{nl}f_{nl} + P_{np}f_{np} + K_{aAP}[AP^-](Y_{1,lyso} + Y_{2,lyso})}{X_{iw}} \right) \frac{X_{lyso}}{X_{ew}} K_{iw^{uu,uu}}^{lyso} K_{cell^{uu,uu}} \quad (28)$$

For mitochondria:

$$K_{mito,u} = \frac{C_{mito}}{C_{media,unbound}} = \left(f_{iw} + \frac{P_{nl}f_{nl} + P_{np}f_{np} + K_{aAP}[AP^-](Y_{1,mito} + Y_{2,mito})}{X_{iw}} \right) \frac{X_{mito}}{X_{ew}} K_{iw^{uu,uu}}^{mito} K_{cell^{uu,uu}} \quad (29)$$

Corresponding concentrations:

$$C_{cell,iw} = K_{cell,iw,u} \cdot C_{media,dissolved,u}$$

$$if f_{iw} = 0, C_{cell,iw} = 0 \quad (30)$$

$$C_{lyso} = K_{lyso,u} \cdot C_{media,dissolved,u}$$

$$if f_{lyso} = 0, C_{lyso} = 0 \quad (31)$$

$$C_{mito} = K_{mito,u} \cdot C_{media,dissolved,u}$$

$$if f_{mito} = 0, C_{mito} = 0 \quad (32)$$

Annexe 2 : Supporting Information (Article modèle PBPK)

Table S1. Test concentrations used in the Fish Embryo Acute Toxicity (FET) test and measured concentration data.

Compound	Nominal water concentration (µM)	Total water concentration (µM)	Total embryo* concentration (µM)
Valproic acid	6.25, 12.5, 25, 50, 100, 200, 400	5, 9, 19, 51, 66, 168, 375**, 238	32, 56, 68, 120, 180, 327, 10387**, 462
2,2-Dimethylvaleric acid	268, 312.5, 322, 386, 463, 500, 556, 667, 800	225, 298, 275, 416, 490, 556, 606, 713, 861	575, 647, 321, 919, 1697, 2148, 2424, 8746, 7053
2-Ethylbutyric acid	100, 200, 400, 800	88, 173, 431, 586	24, 92, 1446, 11142
2-Ethylhexanoic acid	12.5, 25, 50, 100, 200	13, 30, 58, 94, 168	53, 56, 75, 140, 667
2-Methylhexanoic acid	125, 250, 500	176, 320, 638	8, 12, 1775
2-Methylpentanoic acid	177.7, 266.6, 400, 600	287, 372, 614, 762	28, 72, 308, 3734
2-Propylheptanoic acid	12.5, 25, 50	14, 27, 52	32, 53, 52
4-eneValproic acid	200, 400	271, 414	666, 3594
4-Pentenoic acid	414, 500, 538, 700, 910, 1000	480, 481, 539, 775, 840, 860	99, ND***, 222, 97, 6674, 12250
Hexanoic acid	512, 520, 563, 620, 625, 681	544, 559, 552, 637, 500, 666	ND, 28, 92, 6510, 10933, 11382

* Embryo and yolk included.

** Measurements at 72 hpf for a nominal water concentration at 400 µM.

*** ND: below limit of detection.

Table S2. Mass spectrometric and chromatography conditions.

**Mass Spectrometric
Conditions**

Ion Spray Voltage (V)	-4500
Temperature (°C)	50
Gas 1 (psi)	50
Gas 2 (psi)	50
Curtain Gas (psi)	40
CAD Gas (arb)	N/A
Detection	Hi Res
Mass detection range (amu)	50 - 800
Accumulation time (ms)	250

Chromatography Conditions

Column		Phenomenex Kinetex XB-C18 2.6µm 2.1x50mm	
Column Temperature (°C)		60	
Mobile Phase A		0.01% Formic acid (aq)	
Mobile Phase B		MeOH	
Time (min)	Flow Rate (µL/min)	% A	% B
0.00	800	98	2
1.50	800	5	95
1.90	800	5	95
2.00	800	98	2
2.50	800	98	2

Table S3. Physiological parameter values of the zebrafish embryo model.

Parameters	Symbols	Units	Values	References
<i>Initial organ growth time</i>				
Liver	τ_{liver}	minutes	960	Our data
Gut	τ_{gut}	minutes	600	Our data
Skeleton	τ_{skeleton}	minutes	2880	Our data
Eyes	τ_{eyes}	minutes	660	Our data
Brain	τ_{brain}	minutes	540	Our data
Heart	τ_{heart}	minutes	1800	Our data
Skin	τ_{skin}	minutes	1440	Our data
Other tissues	τ_{others}	minutes	0	Our data
<i>Organ growth rate</i>				
Liver	$k_{\text{g, liver}}$	minutes ⁻¹	5.30×10^{-6}	Estimated
Gut	$k_{\text{g, gut}}$	minutes ⁻¹	2.28×10^{-5}	Estimated
Skeleton	$k_{\text{g, skeleton}}$	minutes ⁻¹	1.82×10^{-5}	Estimated
Eyes	$k_{\text{g, eyes}}$	minutes ⁻¹	1.20×10^{-5}	Estimated
Brain	$k_{\text{g, brain}}$	minutes ⁻¹	2.22×10^{-5}	Estimated
Heart	$k_{\text{g, heart}}$	minutes ⁻¹	3.01×10^{-5}	Estimated
Skin	$k_{\text{g, skin}}$	minutes ⁻¹	1.37×10^{-5}	Estimated
Other tissues	$k_{\text{g, others}}$	minutes ⁻¹	3.93×10^{-5}	Estimated
Muscle	$k_{\text{g, muscle}}$	minutes ⁻¹	5.07×10^{-5}	Estimated
<i>Yolk consumption rate</i>	$K_{\text{d, yolk}}$	minutes ⁻¹	3.13×10^{-4}	Estimated
<i>Volume of culture well</i>	V_{well}	L	0.0034	Busquet <i>et al.</i> (2014)
<i>Volume of culture medium</i>	V_{medium}	L	0.001	Busquet <i>et al.</i> (2014)
<i>Volume of liver cell</i>	$V_{\text{cell liver}}$	L	10^{-12}	Our data
<i>Fraction lysosome in cells</i>	f_{lyso}	-	0.01	Our data
<i>Fraction mitochondria in cells</i>	f_{mito}	-	0.1	Our data
<i>Radius of embryo at initial time</i>	$r_{\text{embryo},0}$	mm	0.13	Kimmel <i>et al.</i> (1995a)
<i>Radius of yolk at initial time</i>	$r_{\text{yolk},0}$	mm	0.4	Kimmel <i>et al.</i> (1995a)
<i>Volume of embryo at final time</i>	$V_{\text{embryo (120hpf)}}$	L	3.7×10^{-7}	Estimated
<i>Organ vol. as fractions of total at 120 hpf</i>				
Others	$SC_{\text{others (120hpf)}}$	-	1.99×10^{-1}	Estimated
Muscle	$SC_{\text{muscle (120hpf)}}$	-	2.00×10^{-1}	Estimated

Table S4. Zebrafish embryo-specific partition coefficient values predicted by the VIVD model, for VPA and some analogs.

Compound	Fraction un-ionized in medium	Compartment to water partition coefficients													
		partition coefficients		Medium unbound to compartment partition coefficients											
		Plastic*	Air	Yolk	Liver	Gut	Muscle	Skeleton	Eye	Brain	Heart	Skin	Others	Lysosomes	Mitochondria
	<i>f_{ui}</i>	<i>P_{p:w}</i>	<i>P_{a:w}</i>	<i>P_{mu:yolk}</i>	<i>P_{mu:liver}</i>	<i>P_{mu:gut}</i>	<i>P_{mu:muscle}</i>	<i>P_{mu:os}</i>	<i>P_{mu:eye}</i>	<i>P_{mu:brain}</i>	<i>P_{mu:heart}</i>	<i>P_{mu:skin}</i>	<i>P_{mu:others}</i>	<i>P_{mu:lyso}</i>	<i>P_{mu:mito}</i>
Valproic acid	0.00627	5.20×10 ⁻⁴	2.52	0.374	0.290	0.254	0.289	0.289	0.307	0.318	0.289	0.307	0.231	0.0510	0.486
2,2- Dimethylvaleric acid	0.00990	9.96×10 ⁻⁵	1.84	0.245	0.297	0.272	0.303	0.303	0.332	0.332	0.303	0.332	0.235	0.0170	0.714
2-Ethylbutyric acid	0.00315	2.38×10 ⁻⁴	5.84	0.167	0.207	0.189	0.211	0.211	0.231	0.231	0.211	0.231	0.164	0.0113	0.235
2-Ethylhexanoic acid	0.00627	3.56×10 ⁻⁴	2.52	0.303	0.277	0.246	0.278	0.278	0.300	0.306	0.278	0.300	0.219	0.0349	0.483
2-Methylhexanoic acid	0.0124	1.49×10 ⁻⁴	1.97	0.321	0.328	0.296	0.332	0.332	0.360	0.364	0.332	0.360	0.260	0.0310	0.856
2-Methylpentanoic acid	0.00990	4.56×10 ⁻⁵	1.66	0.208	0.290	0.268	0.297	0.297	0.328	0.326	0.297	0.328	0.229	0.00908	0.712
2-Propylheptanoic acid	0.00627	5.19×10 ⁻³	5.84	2.51	0.708	0.477	0.622	0.622	0.542	0.698	0.622	0.542	0.576	0.548	0.582
4-eneVPA	0.00315	2.38×10 ⁻⁴	5.84	0.167	0.207	0.189	0.211	0.211	0.231	0.231	0.211	0.231	0.164	0.0113	0.235
4-Pentenoic acid	0.00990	1.02×10 ⁻⁵	1.15	0.185	0.286	0.266	0.294	0.294	0.326	0.322	0.294	0.326	0.226	0.00410	0.710
Hexanoic acid	0.00788	8.15×10 ⁻⁵	1.43	0.209	0.274	0.252	0.280	0.280	0.308	0.307	0.280	0.308	0.217	0.0113	0.588

*Unit: dm

Table S5. Physicochemical properties of compounds.

Substance	pKa	MW (g/mol)	logP	Henry's constant (Pa.m ³ /mol)
Valproic acid	4.8	144.2	2.75	0.367
2,2-Dimethylvaleric acid	5.0	130.2	2.01	0.229
2-Ethylbutyric acid	4.8	116.2	1.68	0.184
2-Ethylhexanoic acid	4.8	144.2	2.58	0.367
2-Methylhexanoic acid	5.1	130.2	2.19	0.26
2-Methylpentanoic acid	5.0	116.2	1.66	0.184
2-Propylheptanoic acid	4.8	172.3	3.78	0.731
4-eneVPA	4.5	142.2	2.40	0.731
4-Pentenoic acid	4.9	100.0	0.99	0.0266
Hexanoic acid	4.9	116.2	1.92	0.119

Table S6. Estimation of mean, standard deviation (SD), 95% confidence intervals (IC95%) and maximum posterior (MP) value for the metabolic clearance, and residual uncertainty SD σ , for VPA and its analogs.

Compound	<i>Cl_{met}</i> (L/min)			σ		
	Mean $\pm$ SD	IC 95%	MP	Mean $\pm$ SD	IC 95%	MP
Valproic acid	$3.08 \times 10^{-11} \pm 2.98 \times 10^{-11}$	[1.10×10^{-12} ; 1.10×10^{-10}]	2.26×10^{-13}	4.90 ± 0.728	[3.67;6.48]	4.49
2,2-Dimethylvaleric acid	$2.02 \times 10^{-11} \pm 1.85 \times 10^{-11}$	[7.26×10^{-13} ; 6.86×10^{-11}]	1.66×10^{-13}	4.97 ± 0.755	[3.71;6.62]	4.64
2-Ethylbutyric acid	$6.30 \times 10^{-11} \pm 5.71 \times 10^{-11}$	[1.65×10^{-12} ; 2.17×10^{-10}]	5.93×10^{-13}	4.53 ± 0.809	[3.25;6.22]	4.13
2-Ethylhexanoic acid	$3.37 \times 10^{-11} \pm 3.38 \times 10^{-11}$	[7.58×10^{-13} ; 1.23×10^{-10}]	7.39×10^{-15}	4.19 ± 0.792	[2.97;6.12]	3.59
2-Methylhexanoic acid	$8.24 \times 10^{-11} \pm 6.55 \times 10^{-11}$	[3.19×10^{-12} ; 2.48×10^{-10}]	3.47×10^{-13}	3.65 ± 0.737	[2.49;5.40]	3.18
2-Methylpentanoic acid	$4.61 \times 10^{-11} \pm 4.20 \times 10^{-11}$	[1.25×10^{-12} ; 1.55×10^{-10}]	9.19×10^{-14}	3.54 ± 0.741	[2.36;5.33]	2.95
2-Propylheptanoic acid	$1.86 \times 10^{-11} \pm 1.90 \times 10^{-11}$	[4.73×10^{-13} ; 7.39×10^{-11}]	9.10×10^{-14}	2.86 ± 0.727	[1.85;4.40]	2.25
4-eneVPA	$8.07 \times 10^{-11} \pm 7.34 \times 10^{-11}$	[3.21×10^{-12} ; 2.82×10^{-10}]	1.98×10^{-13}	4.63 ± 0.834	[3.29;6.52]	4.11
4-Pentenoic acid	$4.23 \times 10^{-11} \pm 3.84 \times 10^{-11}$	[1.14×10^{-12} ; 1.46×10^{-10}]	5.56×10^{-13}	4.14 ± 0.769	[2.95;5.96]	3.62
Hexanoic acid	$4.46 \times 10^{-11} \pm 4.18 \times 10^{-11}$	[1.26×10^{-12} ; 1.56×10^{-10}]	1.64×10^{-13}	5.10 ± 0.781	[3.76;6.89]	4.72

Table S7. Estimation of mean, standard deviation (SD), 95% confidence intervals (IC95%) and maximum posterior value for the metabolic clearance, the coefficient partition correction factor, the air to water partition coefficient, and residual fit uncertainty σ , for VPA, its analogs and the 16 compounds studied by Brox *et al.*

	Cl_{met} (L/min)			f_{pc}			$P_{a:w}$			σ		
	Mean $\pm$ SD	IC 95%	MP	Mean $\pm$ SD	IC 95%	MP	Mean $\pm$ SD	IC 95%	MP	Mean $\pm$ SD	IC 95%	MP
Valproic acid	$6.54 \times 10^{-12} \pm 4.89 \times 10^{-12}$	$[3.03 \times 10^{-13}; 1.82 \times 10^{-11}]$	7.80×10^{-14}	7.31 ± 0.919	$[5.66; 9.31]$	6.87	4.10 ± 5.11	$[0.293; 18.7]$	1.03	1.71 ± 0.163	$[1.51; 2.13]$	1.56
2,2- dimethylvaleric acid	$1.78 \times 10^{-10} \pm 1.17 \times 10^{-11}$	$[1.07 \times 10^{-11}; 4.46 \times 10^{-10}]$	8.77×10^{-11}	0.432 ± 0.200	$[0.141; 0.919]$	0.314	2.73 ± 3.28	$[0.210; 12.6]$	0.614	2.16 ± 0.629	$[1.52; 3.84]$	1.53
2-ethylbutyric acid	$5.47 \times 10^{-10} \pm 4.26 \times 10^{-10}$	$[2.25 \times 10^{-11}; 1.55 \times 10^{-9}]$	3.49×10^{-10}	0.259 ± 0.106	$[0.0914; 0.481]$	0.201	9.23 ± 10.3	$[0.672; 39.4]$	2.35	2.77 ± 0.644	$[1.87; 4.40]$	2.26
2-ethylhexanoic acid	$1.43 \times 10^{-10} \pm 1.04 \times 10^{-10}$	$[7.11 \times 10^{-12}; 4.01 \times 10^{-10}]$	6.12×10^{-12}	0.332 ± 0.107	$[0.186; 0.604]$	0.268	3.22 ± 3.72	$[0.248; 14.1]$	0.710	1.75 ± 0.267	$[1.51; 2.48]$	1.53
2-methylhexanoic acid	$2.95 \times 10^{-9} \pm 2.38 \times 10^{-9}$	$[1.66 \times 10^{-10}; 9.15 \times 10^{-9}]$	6.92×10^{-10}	0.0902 ± 0.0832	$[0.0103; 0.332]$	0.033	4.34 ± 8.74	$[0.265; 23.8]$	0.532	3.89 ± 0.797	$[2.58; 5.77]$	3.45
2-methylpentanoic acid	$7.97 \times 10^{-10} \pm 6.91 \times 10^{-10}$	$[3.17 \times 10^{-11}; 2.57 \times 10^{-9}]$	8.18×10^{-12}	0.111 ± 0.0749	$[0.0286; 0.321]$	0.0732	2.87 ± 5.87	$[0.193; 11.9]$	0.441	3.25 ± 0.719	$[2.22; 4.94]$	2.61
2-propylheptanoic acid	$2.63 \times 10^{-10} \pm 2.16 \times 10^{-10}$	$[9.87 \times 10^{-12}; 7.69 \times 10^{-10}]$	1.25×10^{-11}	0.467 ± 0.177	$[0.200; 0.907]$	0.34	8.59 ± 9.64	$[0.629; 33.6]$	1.57	1.97 ± 0.538	$[1.51; 3.33]$	1.51
4-ene-valproic acid	$1.07 \times 10^{-10} \pm 9.17 \times 10^{-11}$	$[4.97 \times 10^{-12}; 3.67 \times 10^{-10}]$	1.76×10^{-11}	1.42 ± 0.913	$[0.346; 4.06]$	0.815	8.98 ± 10.1	$[0.731; 39.0]$	1.50	2.45 ± 0.727	$[1.55; 4.28]$	1.66
4-pentenoic acid	$1.65 \times 10^{-10} \pm 1.55 \times 10^{-10}$	$[8.87 \times 10^{-12}; 6.27 \times 10^{-10}]$	1.63×10^{-11}	0.740 ± 0.433	$[0.150; 1.80]$	0.492	1.94 ± 2.60	$[0.125; 8.58]$	0.353	2.31 ± 0.763	$[1.52; 4.22]$	1.53
Hexanoic acid	$9.56 \times 10^{-10} \pm 8.57 \times 10^{-10}$	$[3.05 \times 10^{-11}; 3.44 \times 10^{-9}]$	3.13×10^{-10}	0.351 ± 0.368	$[0.0286; 1.48]$	0.129	2.67 ± 3.39	$[0.183; 12.4]$	0.490	4.20 ± 0.787	$[2.86; 5.95]$	3.85
2,4-diclorophenoxy-acetic acid	$2.09 \times 10^{-11} \pm 1.80 \times 10^{-11}$	$[6.46 \times 10^{-13}; 6.88 \times 10^{-11}]$	1.27×10^{-12}	10.6 ± 1.84	$[7.57; 15.1]$	9.45	1.73 ± 2.30	$[0.149; 8.92]$	0.310	2.12 ± 0.258	$[1.76; 2.77]$	1.92
Atropine	$1.35 \times 10^{-10} \pm 5.73 \times 10^{-11}$	$[2.28 \times 10^{-11}; 2.44 \times 10^{-10}]$	1.31×10^{-10}	0.0248 ± 0.00405	$[0.0177; 0.0341]$	0.0238	2.17 ± 3.64	$[2.17; 3.64]$	0.507	1.79 ± 0.186	$[1.53; 2.27]$	1.68
Benzocaine	$1.54 \times 10^{-11} \pm 3.48 \times 10^{-12}$	$[8.57 \times 10^{-12}; 2.16 \times 10^{-11}]$	1.57×10^{-11}	2.04 ± 0.522	$[1.43; 3.59]$	1.65	2.29 ± 2.84	$[0.147; 11.3]$	0.322	1.59 ± 0.0798	$[1.50; 1.79]$	1.52
Caffeine	$1.87 \times 10^{-11} \pm 6.14 \times 10^{-12}$	$[4.40 \times 10^{-12}; 2.77 \times 10^{-11}]$	1.84×10^{-11}	3.03 ± 0.605	$[2.09; 4.56]$	2.58	1.92 ± 1.89	$[0.133; 7.26]$	0.339	1.54 ± 0.0487	$[1.50; 1.66]$	1.50
Chloramphenicol	$8.46 \times 10^{-10} \pm 2.84 \times 10^{-10}$	$[2.17 \times 10^{-10}; 1.37 \times 10^{-9}]$	8.41×10^{-10}	0.00164 ± 0.000336	$[0.00111; 0.00244]$	0.00154	1.84 ± 2.41	$[0.125; 8.94]$	0.364	1.96 ± 0.215	$[1.64; 2.45]$	1.77
Cimetidine	$1.72 \times 10^{-10} \pm 1.23 \times 10^{-10}$	$[8.63 \times 10^{-12}; 4.70 \times 10^{-10}]$	6.11×10^{-11}	0.228 ± 0.0742	$[0.129; 0.432]$	0.162	2.26 ± 2.76	$[0.143; 10.5]$	0.298	2.66 ± 0.406	$[2.04; 3.65]$	2.40
Clofibric acid	$1.11 \times 10^{-11} \pm 1.02 \times 10^{-11}$	$[3.00 \times 10^{-13}; 3.81 \times 10^{-11}]$	7.95×10^{-14}	15.1 ± 2.13	$[11.1; 19.2]$	14.3	1.92 ± 2.65	$[0.137; 9.32]$	0.481	2.01 ± 0.285	$[1.63; 2.68]$	1.83
Colchicine	$2.71 \times 10^{-9} \pm 1.06 \times 10^{-9}$	$[4.98 \times 10^{-10}; 4.78 \times 10^{-9}]$	2.38×10^{-9}	0.0149 ± 0.00515	$[0.00755; 0.0284]$	0.0117	2.26 ± 2.73	$[0.145; 10.1]$	0.357	2.51 ± 0.385	$[1.94; 3.46]$	2.24
Cyclophosphamide	$5.35 \times 10^{-10} \pm 1.67 \times 10^{-10}$	$[1.91 \times 10^{-10}; 8.41 \times 10^{-10}]$	5.96×10^{-10}	0.122 ± 0.0419	$[0.0775; 0.256]$	0.0922	2.90 ± 3.70	$[1.50; 1.70]$	0.288	1.55 ± 0.0507	$[1.50; 1.70]$	1.50
Metoprolol	$1.99 \times 10^{-11} \pm 1.10 \times 10^{-11}$	$[1.15 \times 10^{-12}; 4.18 \times 10^{-11}]$	1.83×10^{-11}	0.0466 ± 0.00560	$[0.0353; 0.0575]$	0.0448	1.83 ± 2.35	$[0.141; 8.58]$	0.416	1.55 ± 0.0528	$[1.50; 1.69]$	1.50
Metribuzin	$4.51 \times 10^{-12} \pm 3.78 \times 10^{-12}$	$[1.61 \times 10^{-13}; 1.42 \times 10^{-11}]$	2.53×10^{-13}	0.371 ± 0.0506	$[0.291; 0.486]$	0.316	1.47 ± 1.31	$[0.127; 5.13]$	0.310	1.56 ± 0.0579	$[1.50; 1.72]$	1.50
Phenacetin	$1.63 \times 10^{-11} \pm 8.40 \times 10^{-12}$	$[1.68 \times 10^{-12}; 3.29 \times 10^{-11}]$	1.63×10^{-11}	0.788 ± 0.096	$[0.623; 0.978]$	0.731	1.29 ± 1.06	$[0.131; 4.13]$	0.293	1.54 ± 0.043	$[1.50; 1.66]$	1.50
Phenytoin	$7.70 \times 10^{-12} \pm 6.82 \times 10^{-12}$	$[2.79 \times 10^{-13}; 2.61 \times 10^{-11}]$	5.22×10^{-13}	0.183 ± 0.0539	$[0.124; 0.352]$	0.140	2.50 ± 3.25	$[0.159; 12.0]$	0.316	1.94 ± 0.184	$[1.66; 2.37]$	1.81
Sulfamethoxazole	$8.13 \times 10^{-11} \pm 4.56 \times 10^{-11}$	$[4.84 \times 10^{-12}; 1.78 \times 10^{-10}]$	7.20×10^{-11}	6.31 ± 1.04	$[4.59; 8.55]$	5.98	1.73 ± 2.00	$[0.160; 7.88]$	0.317	1.79 ± 0.176	$[1.53; 2.19]$	1.64
Theophylline	$1.90 \times 10^{-11} \pm 1.00 \times 10^{-11}$	$[1.35 \times 10^{-12}; 3.89 \times 10^{-11}]$	1.33×10^{-11}	2.97 ± 1.08	$[1.70; 6.06]$	2.17	2.42 ± 3.31	$[0.135; 12.9]$	0.394	1.95 ± 0.262	$[1.58; 2.59]$	1.76
Thiacloprid	$7.76 \times 10^{-12} \pm 7.00 \times 10^{-12}$	$[2.26 \times 10^{-13}; 2.61 \times 10^{-11}]$	1.05×10^{-13}	0.196 ± 0.081	$[0.125; 0.438]$	0.138	3.14 ± 4.72	$[0.124; 20.0]$	0.253	1.74 ± 0.163	$[1.52; 2.14]$	1.57

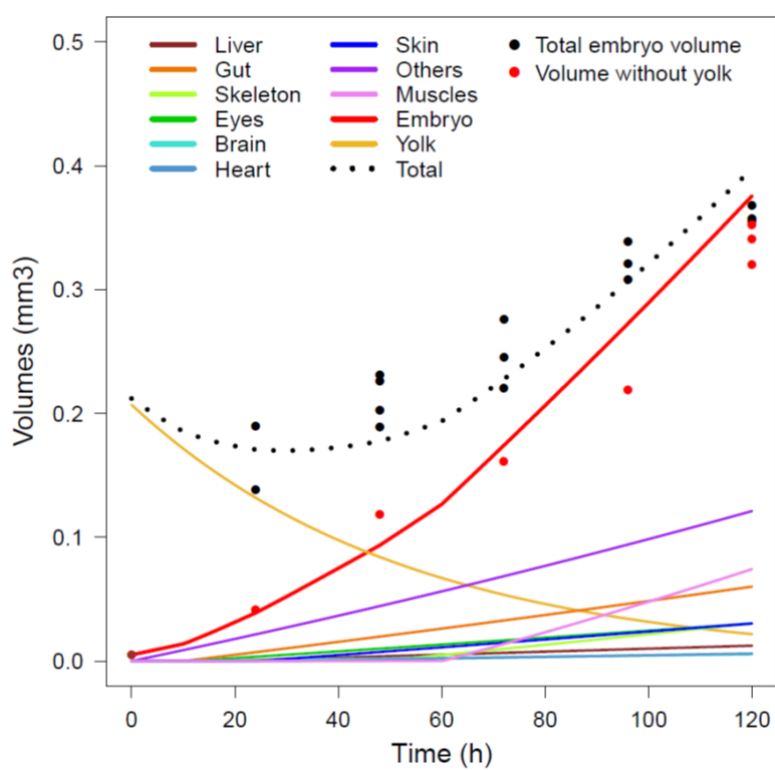


Figure S1. Observed (points) and modeled (lines) time course of the zebrafish embryo's total volume and individual organ volumes.

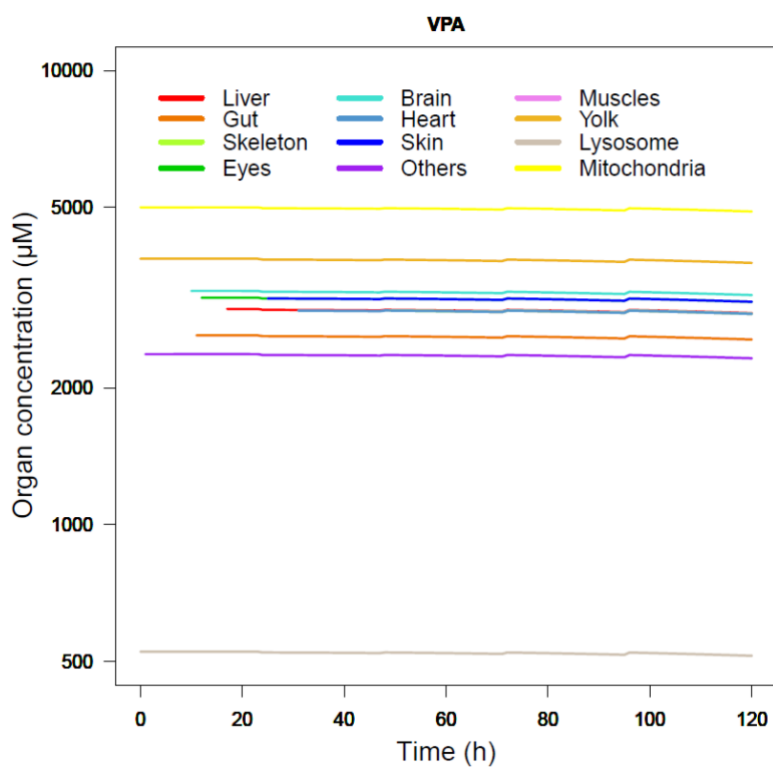


Figure S2. Organ concentrations (μM) as a function of time (h), after estimating Cl_{met} and f_{pc} for VPA.

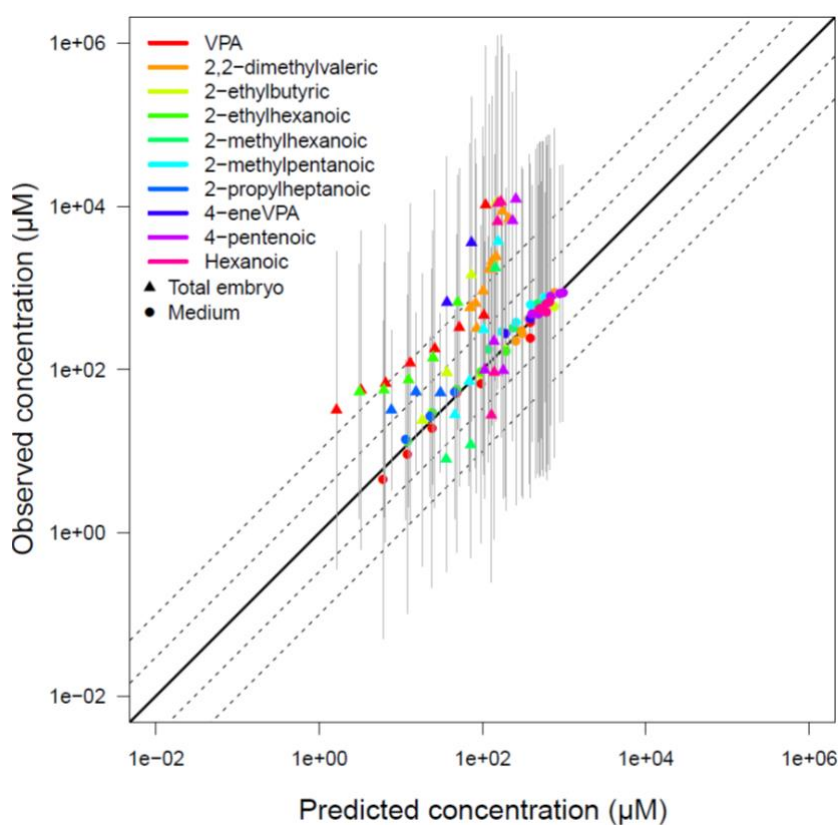


Figure S3. Observed versus predicted concentrations of VPA and nine analogs, in the zebrafish embryo and culture medium. This is the best fit obtained after Cl_{met} estimation by data fitting. The black line corresponds to perfect fit. Dashed lines correspond to the three- and ten-fold error intervals. The grey bars correspond to $\pm$ one residual SD (σ).

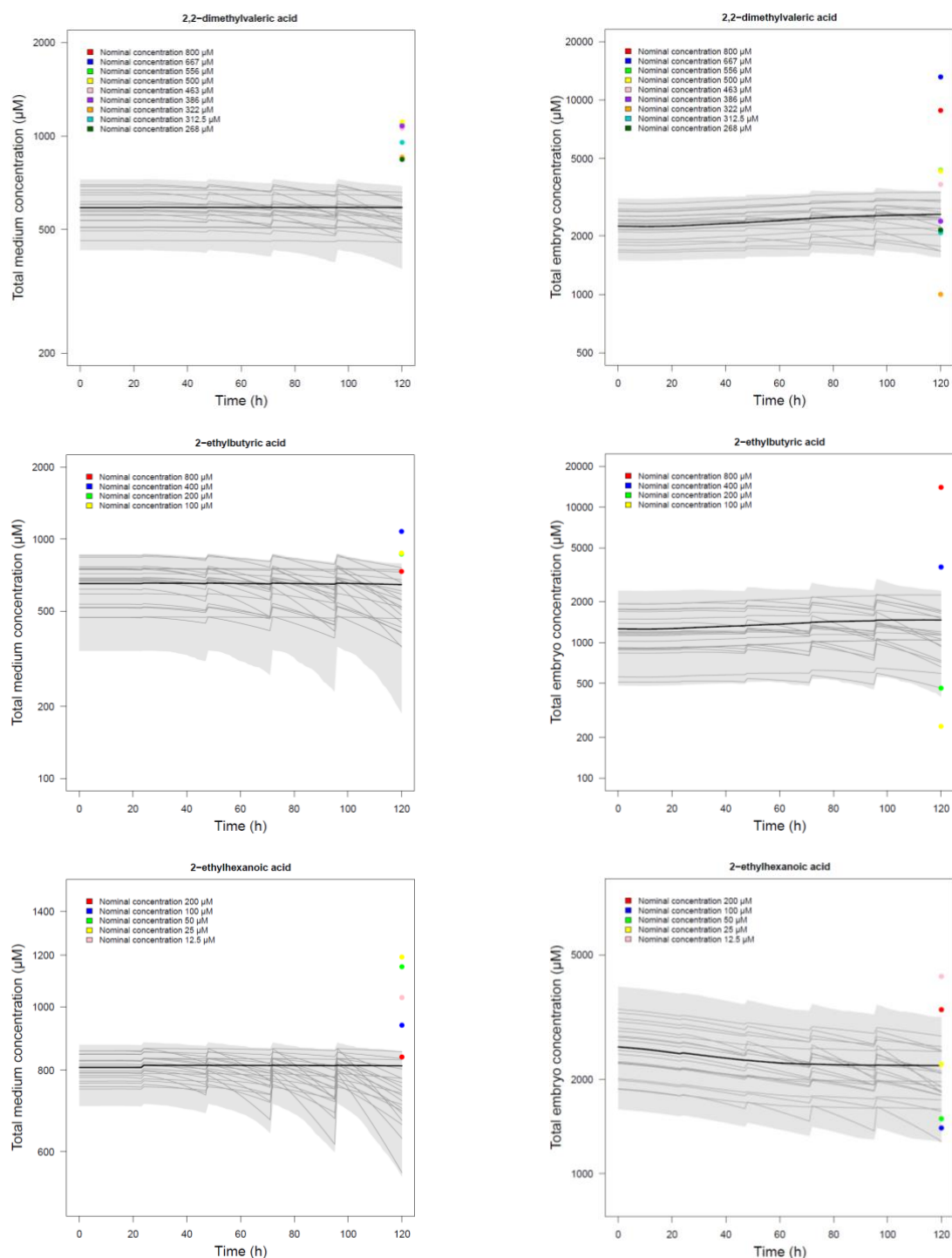


Figure S4. Predicted (lines) and observed (points) VPA analogs' concentrations in medium (left) and in total embryo (right) as a function of time, after estimating Cl_{met} and f_{pc} . All concentrations were normalized to a nominal dose of 1mM. The grey area defines the 95% confidence interval. The thick black line is the maximum posterior predicted concentration time-course. The thin lines are 20 predictions obtained using random parameter vectors drawn from their posterior distribution.

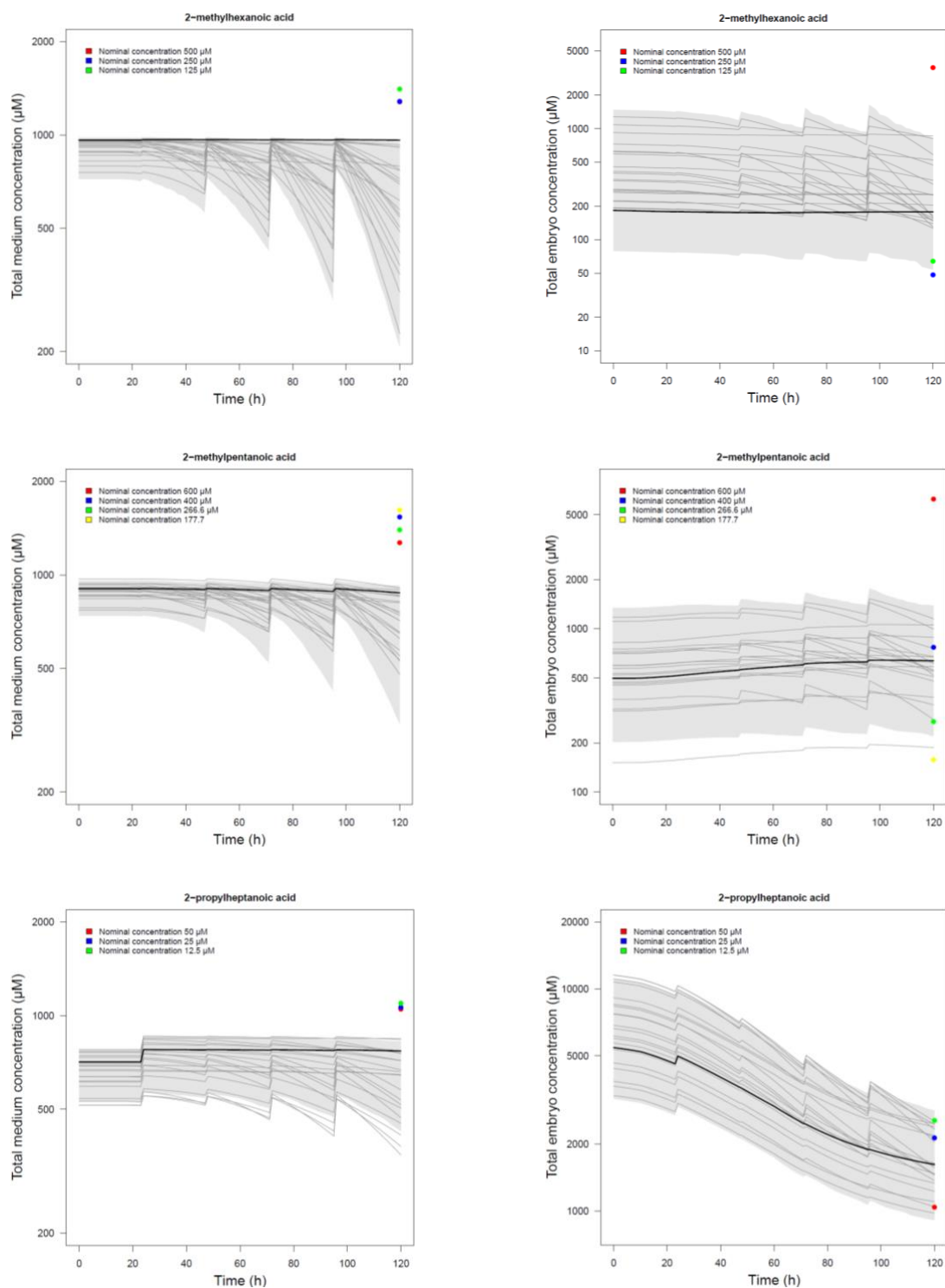


Figure S4 (followed).

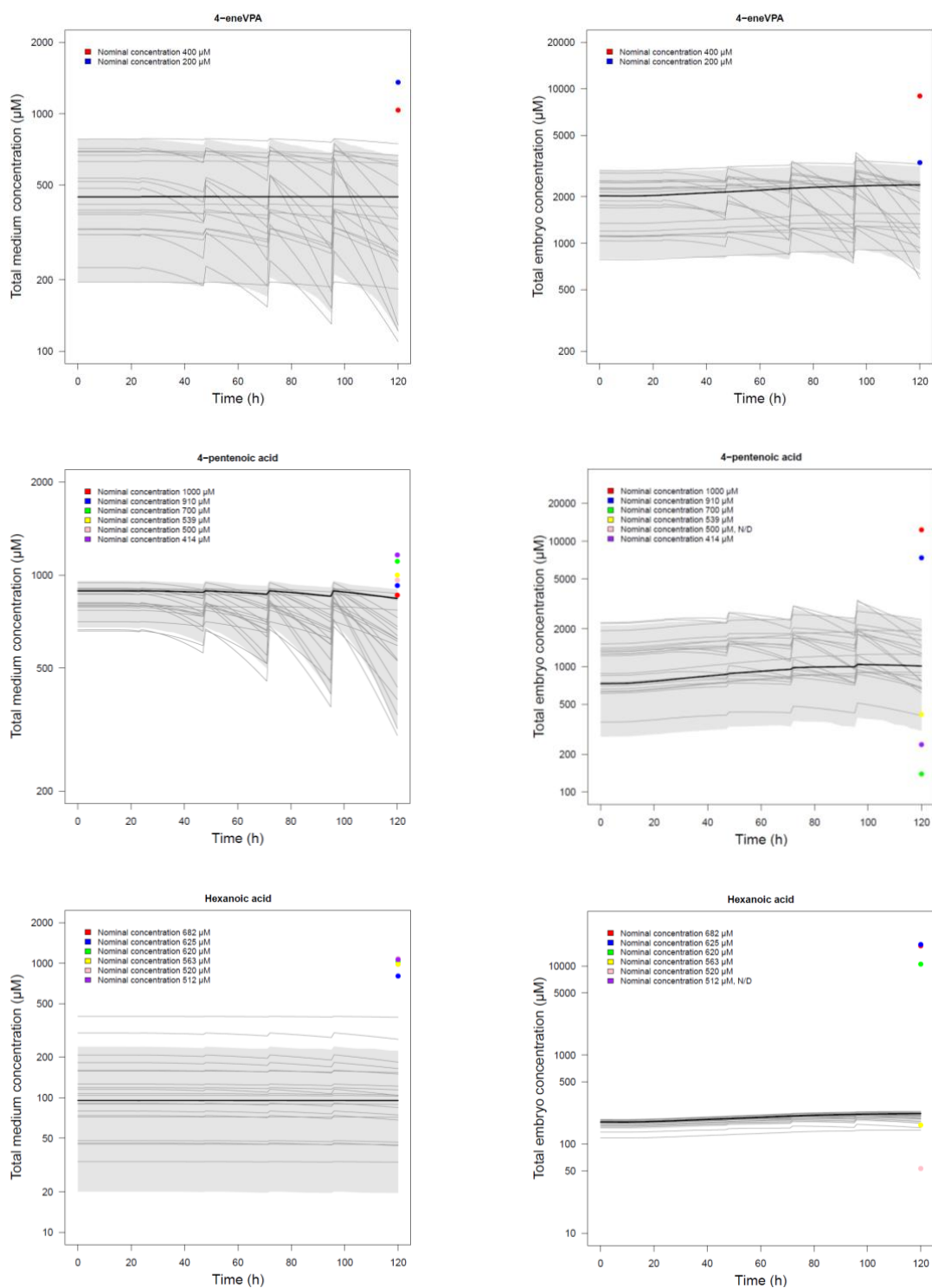


Figure S4 (followed).

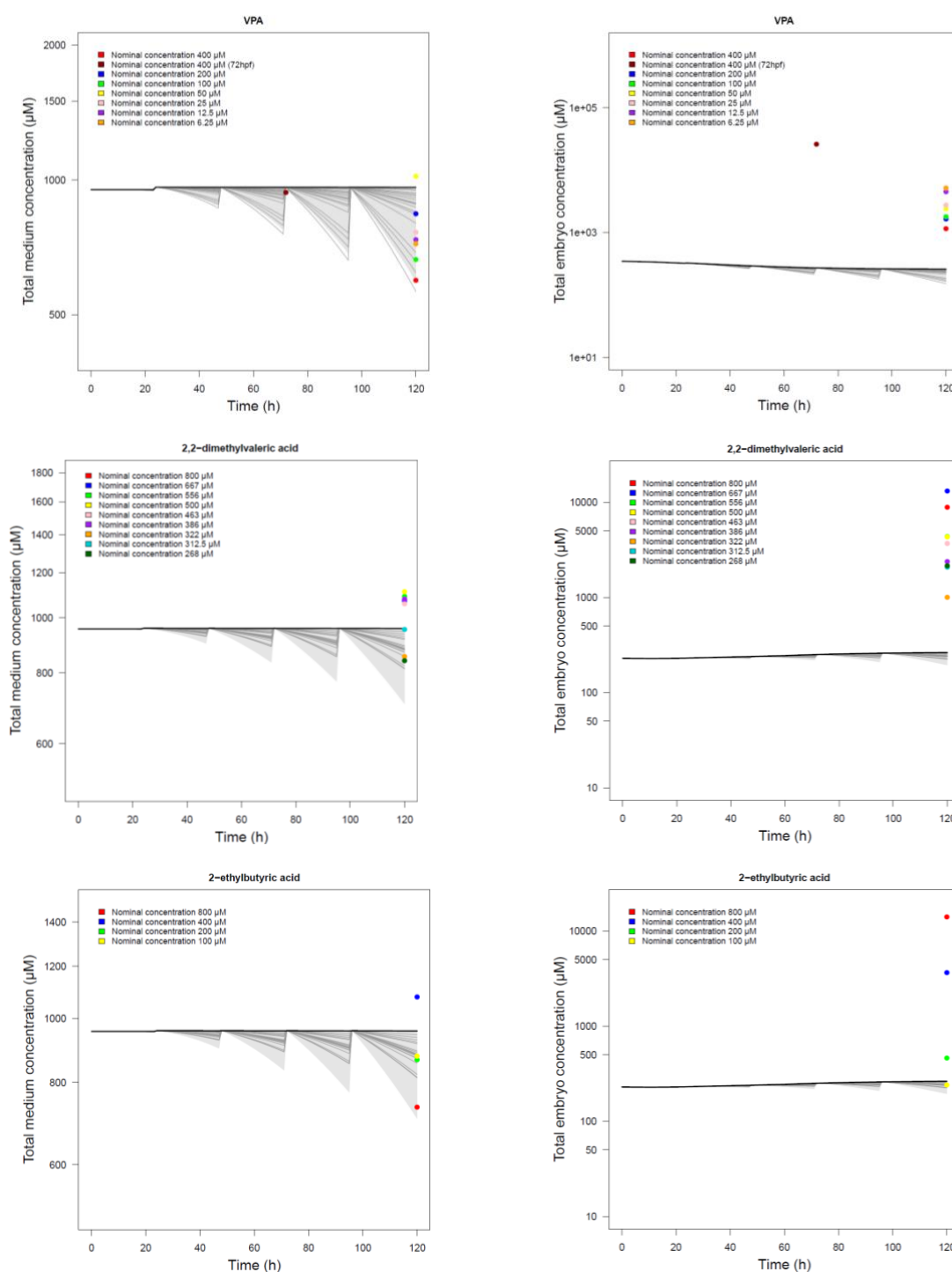


Figure S5. Predicted (lines) and observed (points) VPA and analogs' concentrations in medium (left) and in total embryo (right) as a function of time, after estimating Cl_{met} only. All concentrations were normalized to a nominal dose of 1mM. The grey area defines the 95% confidence interval. The thick black line is the maximum posterior predicted concentration time-course. The thin lines are 20 predictions obtained using random parameter vectors drawn from their posterior distribution.

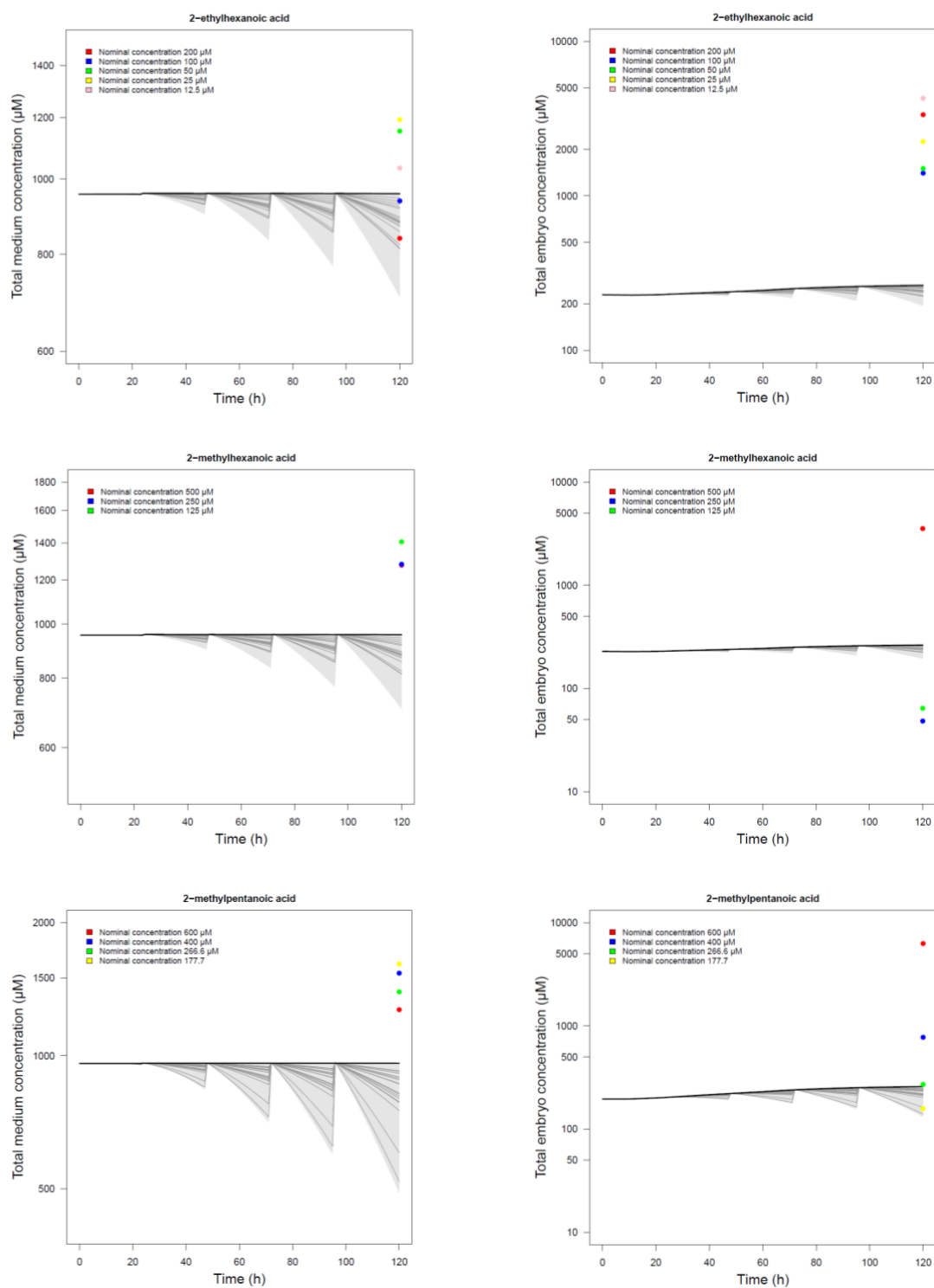


Figure S5 (followed).

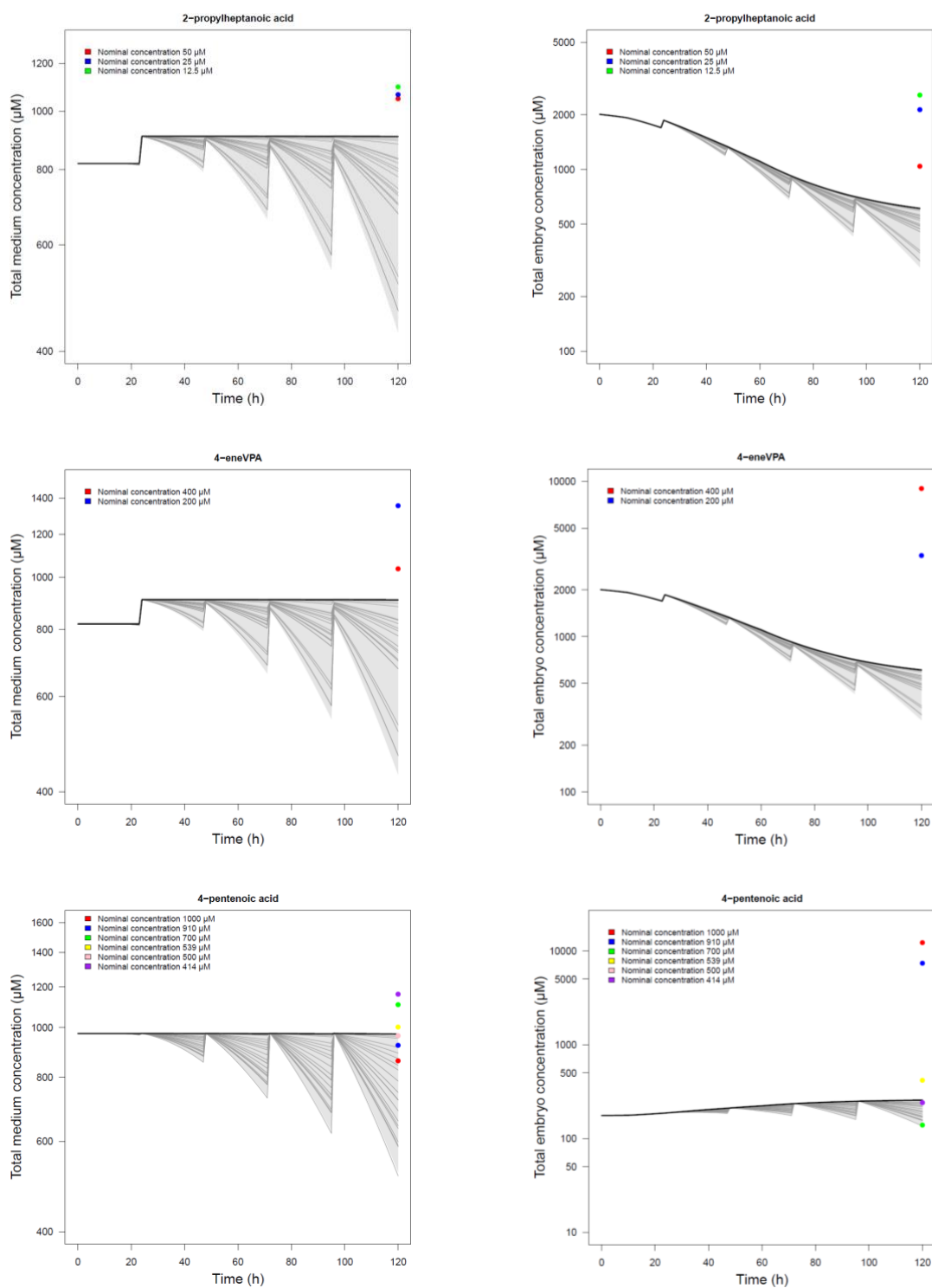


Figure S5 (followed).

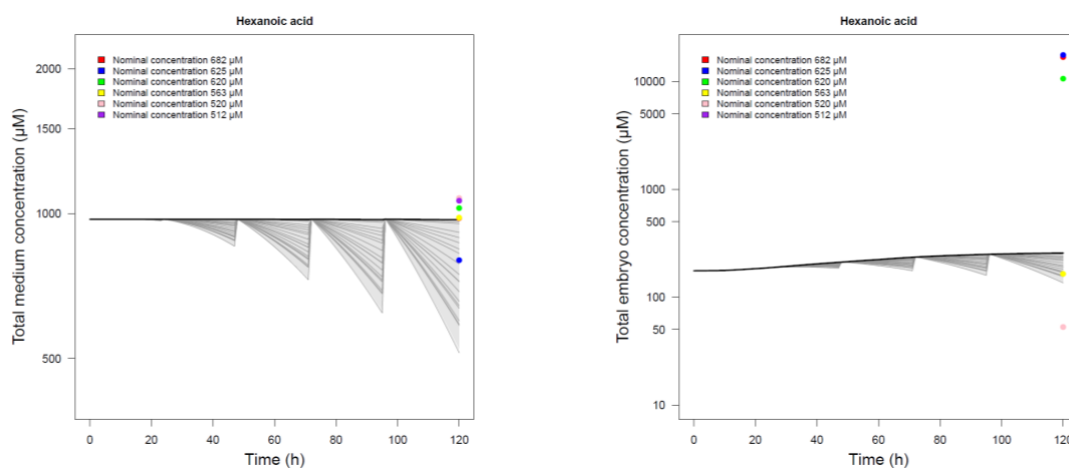


Figure S5 (followed).

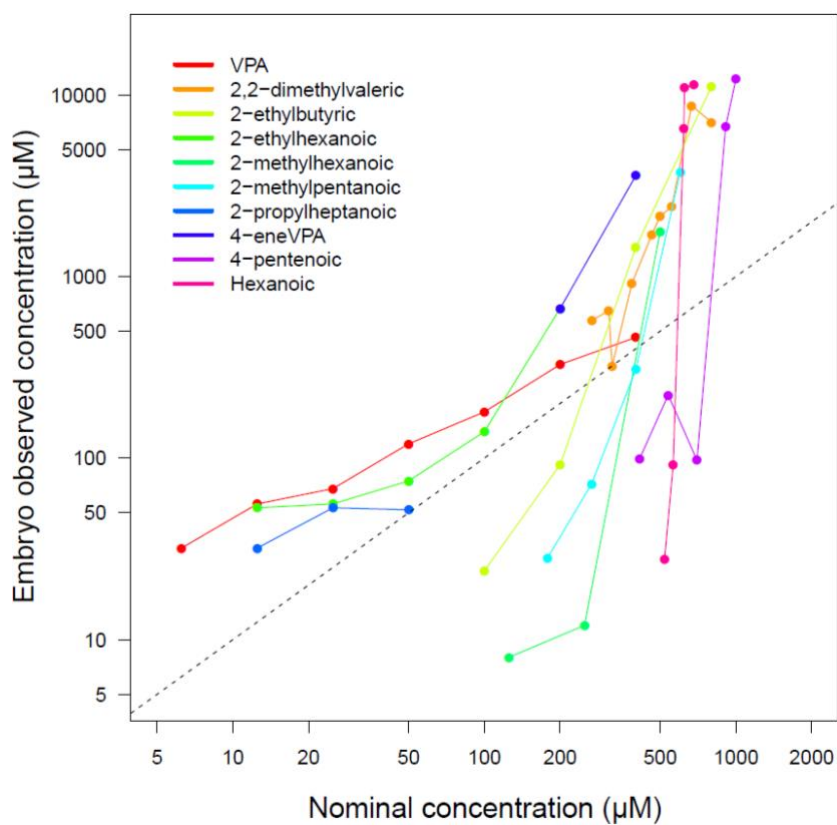


Figure S6. VPA and analogs observed concentrations in total embryo as a function of nominal concentration.

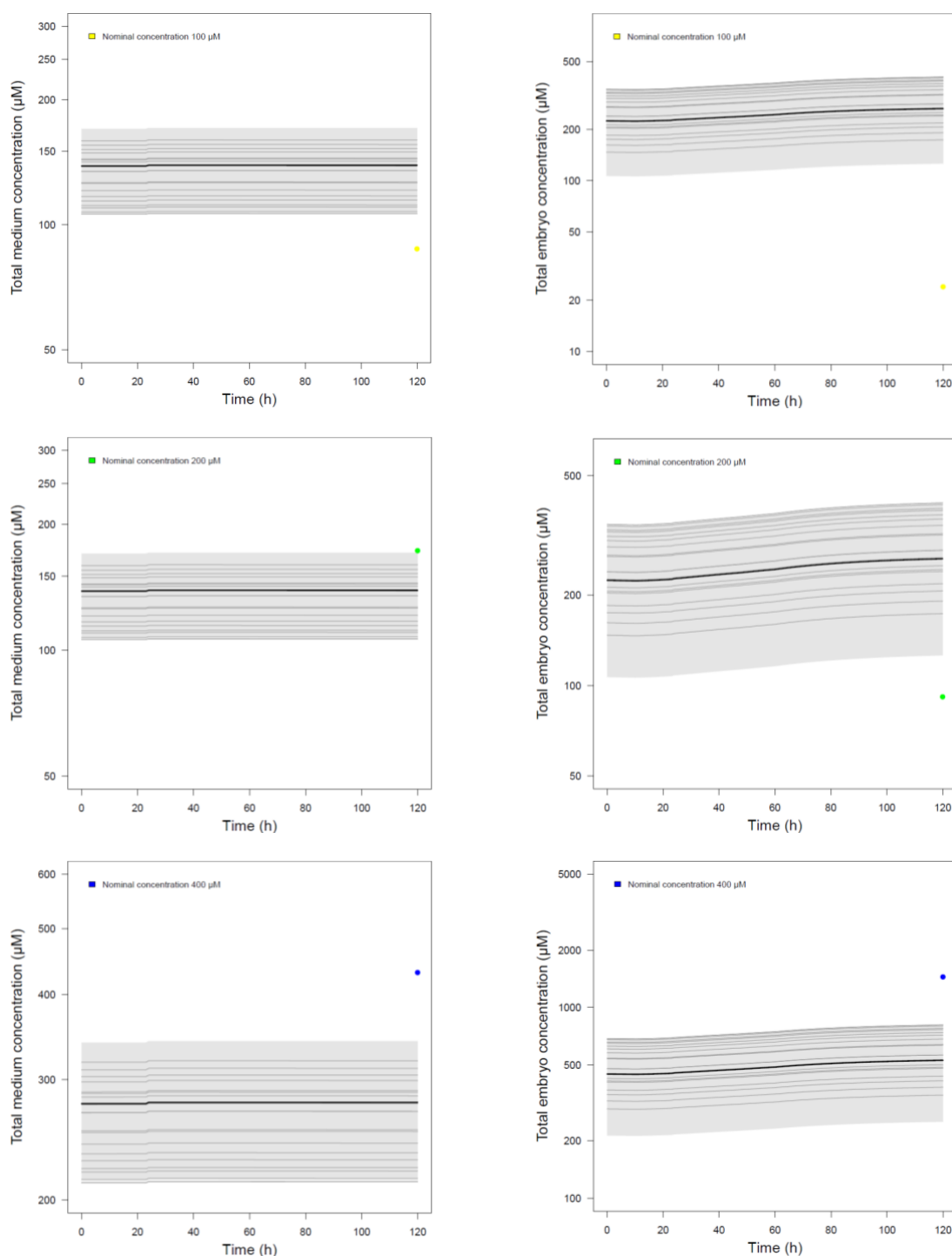


Figure S7. Predicted (lines) and observed (points) 2-ethylbutyric acid concentrations in medium (left) and in total embryo (right) as a function of time, after estimating V_{max} , K_m and f_{pc} (*i.e.*, using saturable Michaelis-Menten metabolism). Linearity is not guaranteed here, and there is one plot per nominal concentrations. The grey areas define the 95% confidence interval. The thick black lines are the maximum posterior predicted concentration time-courses. The thin lines are 20 predictions obtained using random parameter vectors drawn from their posterior distribution.

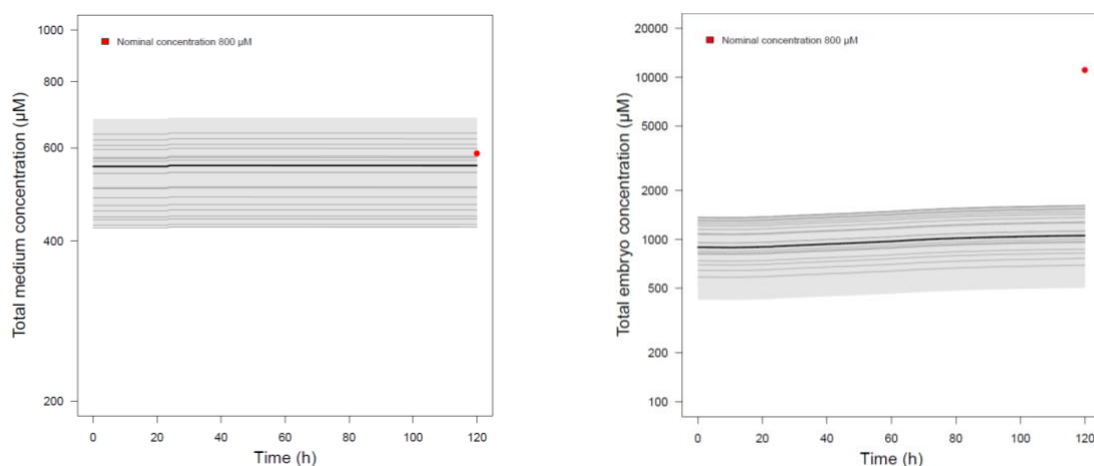


Figure S7 (followed).

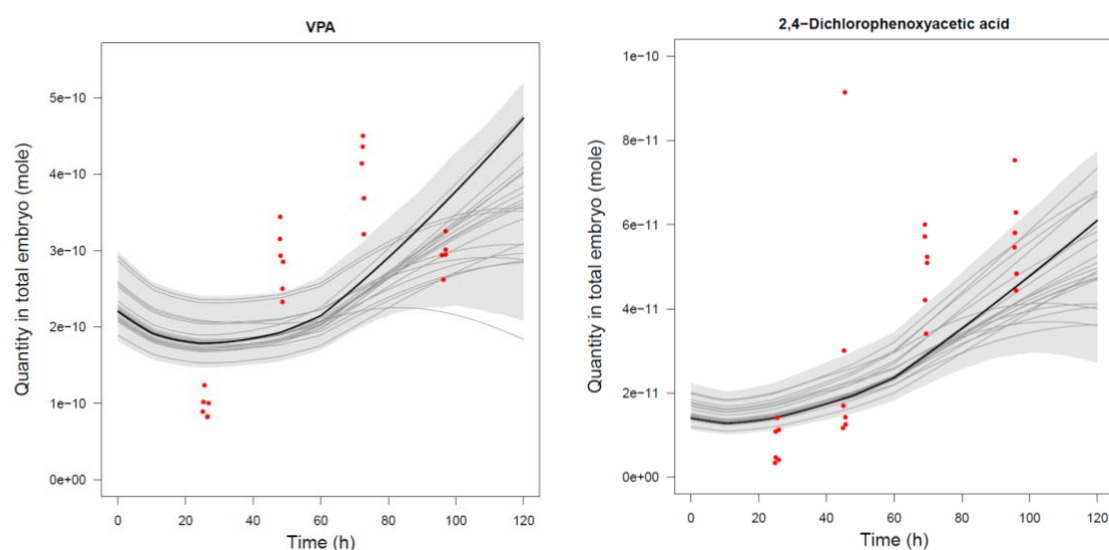


Figure S8. Predicted (lines) and observed (points) total quantities in embryo as a function of time for the 17 compounds studied by Brox *et al.* The grey area defines 95% confident interval. Bold black line is the maximum posterior predicted time-course quantities in the embryo and the other ones are 20 predicted quantity-time courses determined from different estimate parameter values randomly chosen from 1,000 iterations performed. The predictions are obtained with a simultaneous estimation of Cl_{met} , f_{pc} and $P_{a:w}$.

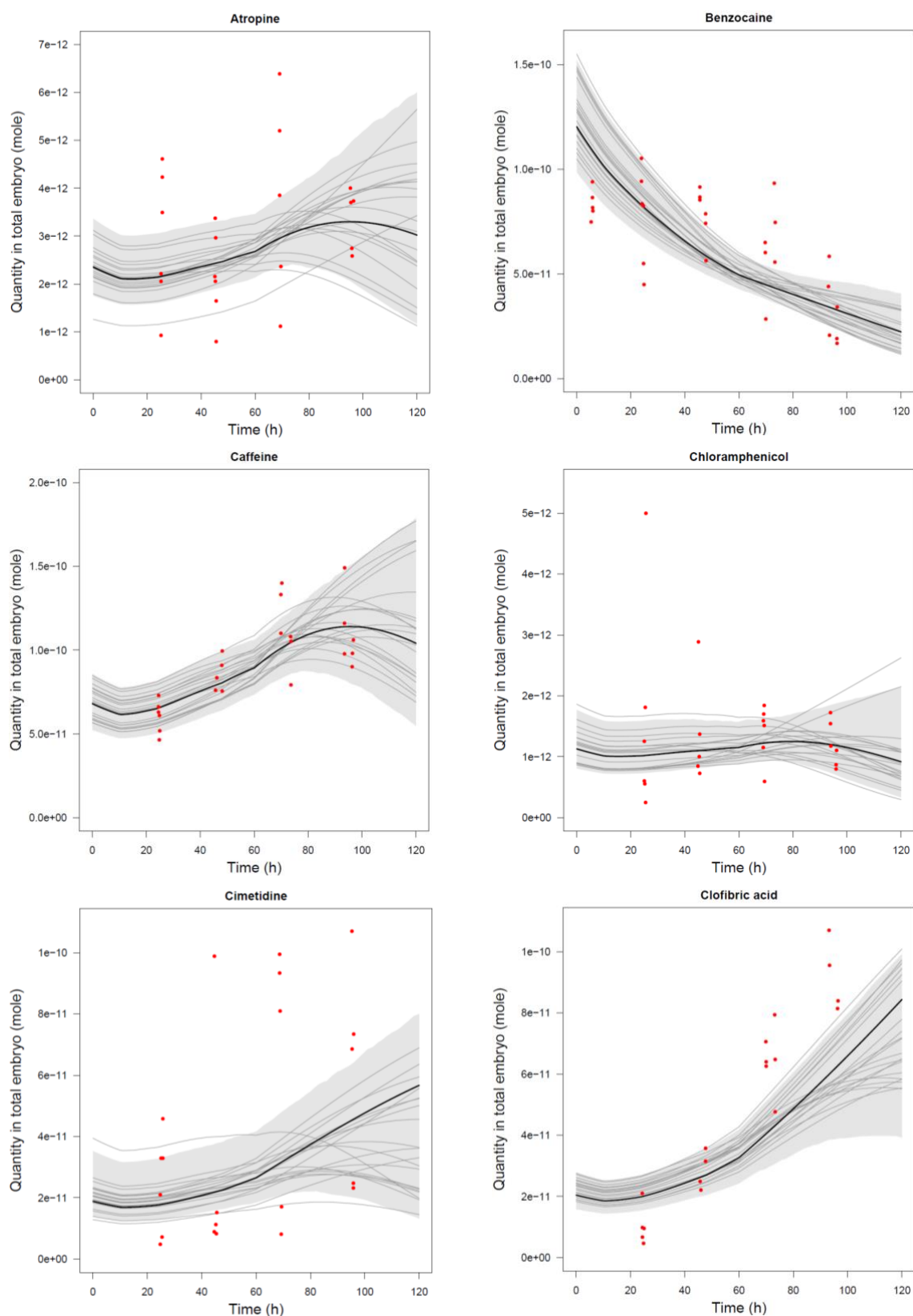


Figure S8 (followed).

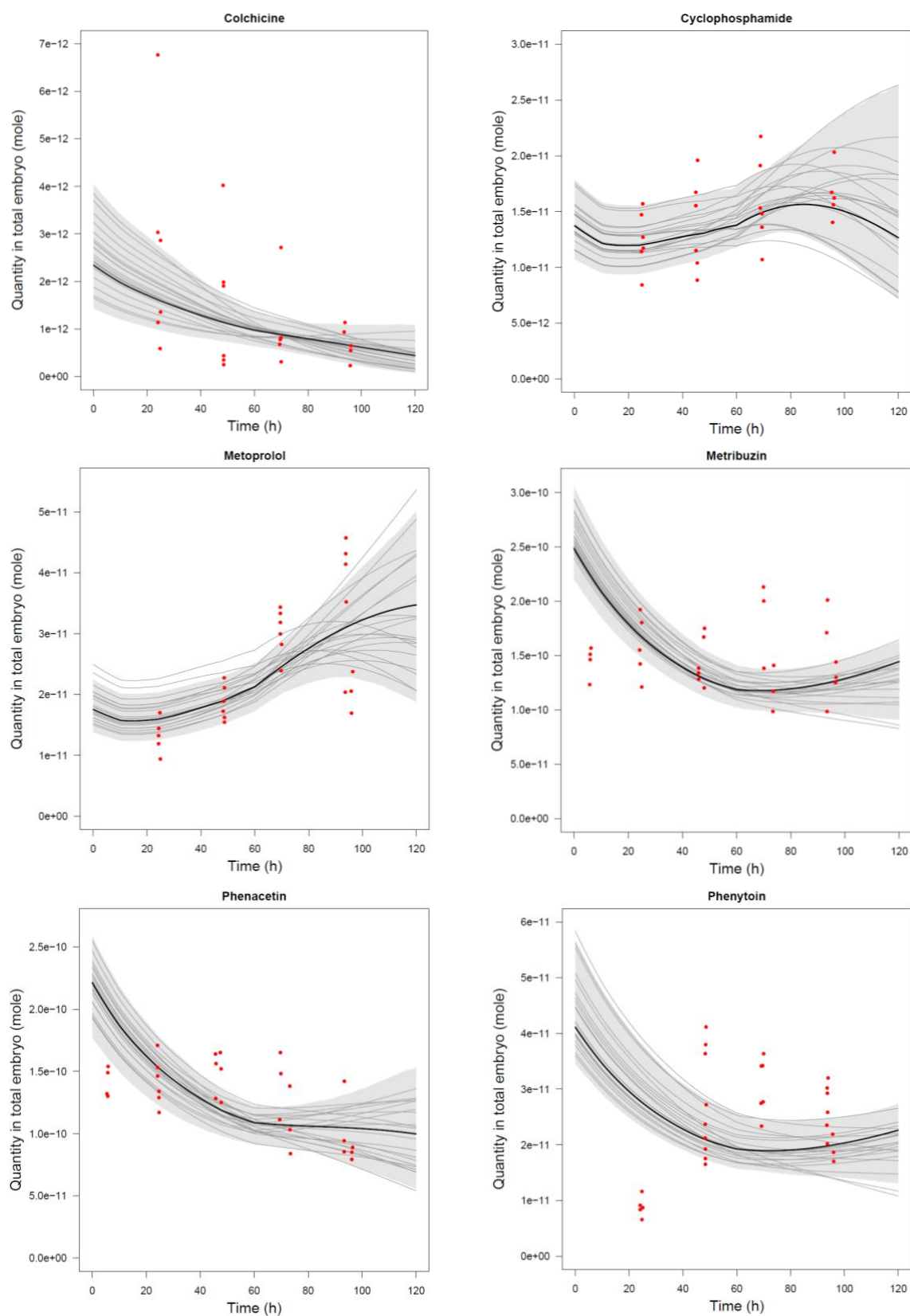


Figure S8 (followed).

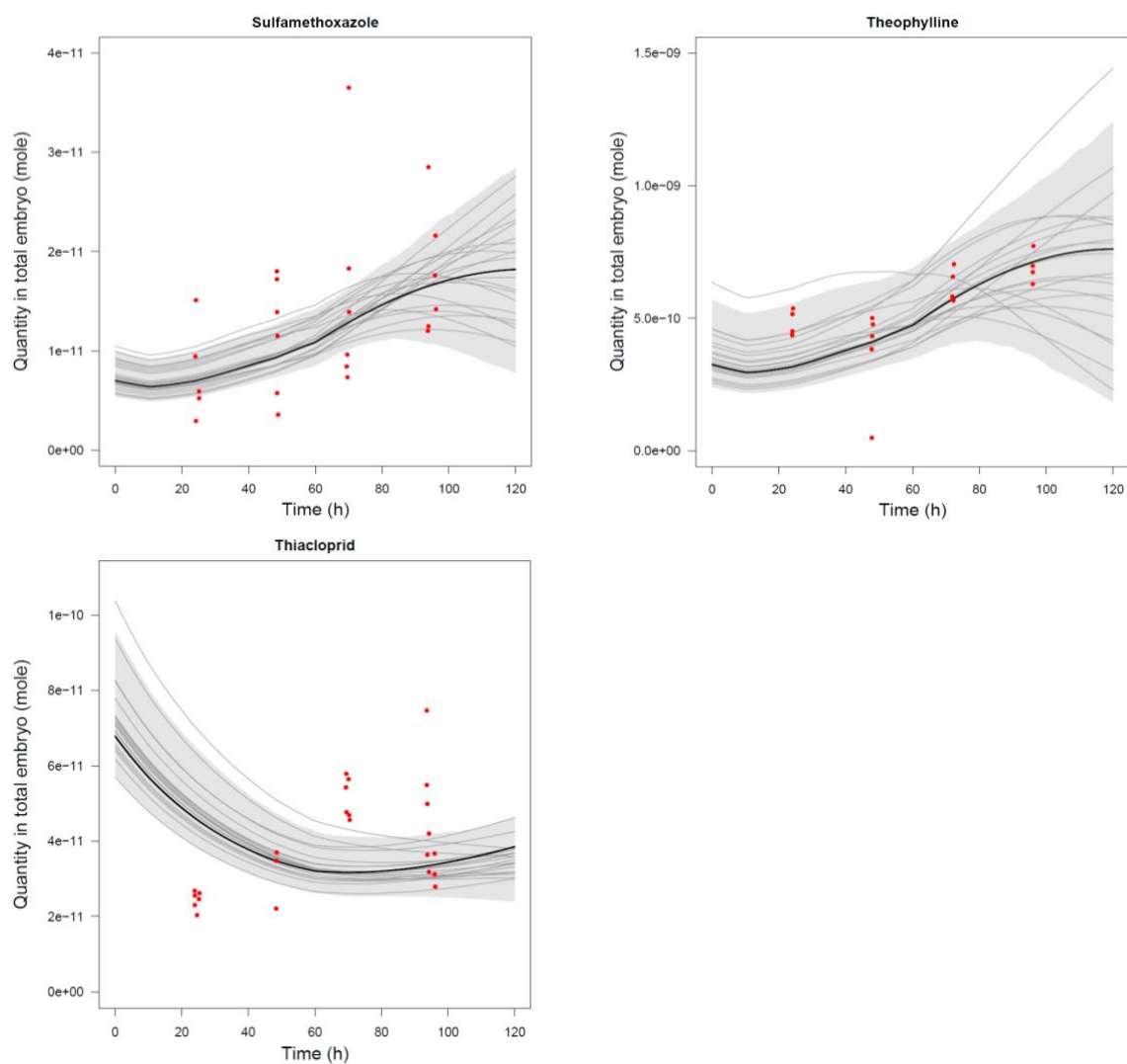


Figure S8 (followed).

Model code

```

States = {Q_labile,      # Sum of Q in medium and air, per embryo (moles)
          Q_fixed,      # Sum of Q in plastic and embryo, per embryo (moles)
          Q_met};      # Q of metabolites in system, per embryo (moles)

Inputs = {Event_labile, Event_fixed}; # Dummy variables for discontinuities

Outputs = {C_medium_u,  # Concentration of free parent in medium (mol/L)
           C_medium,    # ~ of parent in medium
           C_yolk,      # ~ in yolk
           C_liver,     # ~ in liver (metabolizes)
           C_gut,       # ~ in digestive tract
           C_muscle,    # ~ in muscles
           C_skeleton,  # ~ in skeleton
           C_eyes,      # ~ in eyes
           C_brain,     # ~ in brain
           C_heart,     # ~ in heart
           C_skin,      # ~ in skin pigmented cells
           C_others,    # ~ in other tissues
           C_embryo,    # ~ in embryo, not including yolk
           C_embryo_total, # ~ in embryo, including yolk

           C_lyso,      # ~ in lysosomes
           C_mito,      # ~ in mitochondria

           C_air,       # ~ in air
           C_plastic,   # ~ on plastic (mol/dm2)
           Q_check,     # ~ mass balance check, per embryo

           Q_medium,    # Quantity per embryo in medium (mol)
           Q_yolk,      # ~ in yolk
           Q_liver,     # ~ in liver
           Q_gut,       # ~ in digestive tract
           Q_muscle,    # ~ in muscles
           Q_skeleton,  # ~ in skeleton
           Q_eyes,      # ~ in eyes
           Q_brain,     # ~ in brain
           Q_heart,     # ~ in heart
           Q_skin,      # ~ in skin pigmented cells
           Q_others,    # ~ in other tissues
           Q_embryo,    # ~ in embryo, not including yolk
           Q_embryo_total, # ~ in embryo, including yolk

           Q_lyso,      # ~ in lysosomes, per embryo
           Q_mito,      # ~ in mitochondria, per embryo

           Q_air,       # ~ in air, per embryo
           Q_plastic,   # ~ bound on plastic, per embryo
           Q_parent,    # sum of Q_labile and Q_fixed, per embryo

           L_embryo,    # length of the embryo (dm)

           V_yolk,      # volume per embryo of yolk (L)
           V_embryo,    # ~ embryo
           V_embryo_total, # ~ total embryo
           V_liver,     # ~ liver
           V_gut,       # ~ gut
           V_muscle,    # ~ muscles
           V_skeleton,  # ~ skeleton
           V_eyes,      # ~ eyes
           V_brain,     # ~ brain
           V_heart,     # ~ heart
           V_skin,      # ~ skin
           V_others,    # ~ others embryonic tissues
           V_air,       # ~ air in the system
           f_labile,    # Labile fraction (in medium and air) of Q_parent

```

```

S_medium,      # surface area of medium in contact with wall (dm2)
V_app,         # Total apparent volume of distribution per embr. (L)
V_app_e);      # Embryo apparent volume of distribution (L)

# =====

# Total dose of parent (mol) in total medium at start
Dose;

# Number of embryos per well
N_embryo = 1;

# Volume of a culture well (L)
V_well;

# Volume of culture medium at time 0 (L)
V_medium;

# Starting times for organ growth (minutes) (from K. Brotzmann)
tau_liver      = 16 * 60;
tau_gut        = 10 * 60;
tau_skeleton   = 48 * 60;
tau_eyes       = 11 * 60;
tau_brain      = 9 * 60;
tau_heart      = 30 * 60;
tau_skin       = 24 * 60;
tau_others     = 0 * 60; # starts at zero
tau_muscle     = 60 * 60;

# Organ growth rates (1/min), FB adjustment from rescaled organ volumes as
# fractions of total embryo (without yolk) at 120 hpf
K_g_liver      = 5.29509e-06;
K_g_gut        = 2.2779e-05;
K_g_skeleton   = 1.82294e-05;
K_g_eyes       = 1.20414e-05;
K_g_brain      = 2.22253e-06;
K_g_heart      = 3.01003e-06;
K_g_skin       = 1.3672e-05;
K_g_others     = 3.92738e-05;
K_g_muscle     = 5.07074e-05;

# Yolk consumption rate constant (1/min)
K_d_yolk = 0.000313;

# Yolk volume at time zero (L)
V_yolk_0 = 0.207E-6;

# Volume of the embryo without yolk at 0 and 120 hpf (L)
V_embryo_0    = 0.005E-6;
V_embryo_120  = 3.70597E-07;

# Flag for Michaelis-Menten vs linear metabolism (0 = linear)
Michaelis = 0;

# Metabolic clearance per liver cell, if linear metabolism, L/min
K_met;

# Maximum formation rate of metabolite, if saturable metabolism, mol/min
Vmax;

# Michaelis-Menten constant, if saturable metabolism, mol/L
Km;

# Diameter of a well (dm)
D_well;

# Number Pi

```

```

Pi = 3.14159265358979323846;

# Preincubation to saturate plastic binding (yes=1, no=0)
preincubation = 0;

# Volume of a liver cell (L)
V_cell_liver = 1e-12;

# Fraction (v/v) lysosome in cells
f_lyso = 0.01;

# Fraction (v/v) mitochondria in cells
f_mito = 0.1;

# Fraction unbound in medium, corrected for serum dilution
fu_diluted;

# Fraction unionized in medium
fui;

# Plastic / water partition coefficient (dm)
P_pw;

# Fudge factor for partition coefficients
ffpc = 1;

# half-time for equilibration between Q_fixed and Q_labile: 1 second
k_diff = 41.59; # log(2) * 60;

# Statistical parameter
sigma;

# -----
# Parameters recomputed in initialize section
# Must be specified in the input file

# Air / water partition coefficient
P_aw;

# Medium unbound / lysosome partition coefficient
P_mu_ly;

# Medium unbound / mitochondria partition coefficient
P_mu_mi;

# Medium unbound / yolk partition coefficient
P_mu_yolk;

# Medium unbound / liver partition coefficient
P_mu_liver;

# Idem for other organs
P_mu_gut;
P_mu_muscle;
P_mu_skeleton;
P_mu_eye;
P_mu_brain;
P_mu_heart;
P_mu_skin;
P_mu_others;

# -----
# Parameters computed in initialize section and eventually updated in Dynamics

# Volume of well per embryo (L)
V_well_embryo;

# Volume of medium per embryo (L)

```

```

V_medium_embryo;

# -----
Initialize { # assumes that start time is zero, should be revised if not

    P_mu_yolk      = P_mu_yolk      * ffpc;
    P_mu_liver     = P_mu_liver     * ffpc;
    P_mu_gut       = P_mu_gut       * ffpc;
    P_mu_muscle    = P_mu_muscle    * ffpc;
    P_mu_skeleton  = P_mu_skeleton  * ffpc;
    P_mu_eye       = P_mu_eye       * ffpc;
    P_mu_brain     = P_mu_brain     * ffpc;
    P_mu_heart     = P_mu_heart     * ffpc;
    P_mu_skin      = P_mu_skin      * ffpc;
    P_mu_others    = P_mu_others    * ffpc;
    P_mu_ly        = P_mu_ly        * ffpc;
    P_mu_mi        = P_mu_mi        * ffpc;
    P_aw           = P_aw           * ffpc;

    # Compute volumes, surfaces, quantities PER EMBRYO
    V_well_embryo  = V_well / N_embryo;
    V_medium_embryo = V_medium / N_embryo;

    # Surface area of medium in contact with plastic per embryo, but
    # without embryos in the well (dm2)
    S_medium_0     = 4 * V_medium_embryo / D_well + Pi * 0.25 * D_well * D_well;

    # Concentration unbound in medium without embryo in the well (M)
    C_medium_u_0   = Dose / (V_medium / fu_diluted);

    # "Concentration" (quantity per unit surface area) on plastic at start
    # turned on/off by the preincubation switch
    C_plastic_0    = preincubation * P_pw * C_medium_u_0;

    # Quantity on plastic at start, per embryo
    Q_plastic_0    = C_plastic_0 * S_medium_0;

    # Total quantity of parent chemical per embryo (mol)
    Q_parent       = (Dose / N_embryo) + Q_plastic_0;

    # Total volume of one embryo at start (L)
    V_embryo_total = V_yolk_0 + V_embryo_0;

    # Total volume of medium + embryo, for one embryo (L)
    V_content      = V_medium_embryo + V_embryo_total;

    # Surface area of medium in contact with plastic per embryo,
    # with the embryo in the well (dm2)
    S_medium       = 4 * V_content / D_well + Pi * 0.25 * D_well * D_well;

    # Volume of air (L) in the (sealed) well head-space, per embryo
    V_air          = V_well_embryo - V_content;

    # Total apparent volume of distribution per embryo (L)
    V_app = V_medium_embryo / fu_diluted +
            P_aw * fui * V_air +
            P_pw * S_medium +
            P_mu_yolk * V_yolk_0 +
            P_mu_others * V_embryo_0;

    # Labile fraction (in medium and air) of Q_parent (is a ratio of volumes)
    f_labile       = (V_medium_embryo / fu_diluted +
                    P_aw * fui * V_air) / V_app;

    # Quantities per embryo (mol)
    Q_labile       = f_labile * Q_parent;
    Q_fixed        = (1 - f_labile) * Q_parent;
}

```

```

Dynamics {

  # Scaling coefficients of organ volumes (change with time)
  sc_liver    = (t < tau_liver    ?
                 0 : exp(K_g_liver * (t - tau_liver    )) - 1);
  sc_gut      = (t < tau_gut      ?
                 0 : exp(K_g_gut   * (t - tau_gut      )) - 1);
  sc_skeleton = (t < tau_skeleton ?
                 0 : exp(K_g_skeleton * (t - tau_skeleton)) - 1);
  sc_eyes     = (t < tau_eyes     ?
                 0 : exp(K_g_eyes  * (t - tau_eyes     )) - 1);
  sc_brain    = (t < tau_brain    ?
                 0 : exp(K_g_brain  * (t - tau_brain    )) - 1);
  sc_heart    = (t < tau_heart    ?
                 0 : exp(K_g_heart  * (t - tau_heart    )) - 1);
  sc_skin     = (t < tau_skin     ?
                 0 : exp(K_g_skin   * (t - tau_skin     )) - 1);
  sc_others   = (t < tau_others   ?
                 0 : exp(K_g_others * (t - tau_others   )) - 1);
  sc_muscle   = (t < tau_muscle   ?
                 0 : exp(K_g_muscle * (t - tau_muscle   )) - 1);

  # Organ volumes (L):
  V_liver     = V_embryo_120 * sc_liver;
  V_gut       = V_embryo_120 * sc_gut;
  V_skeleton  = V_embryo_120 * sc_skeleton;
  V_eyes      = V_embryo_120 * sc_eyes;
  V_brain     = V_embryo_120 * sc_brain;
  V_heart     = V_embryo_120 * sc_heart;
  V_skin      = V_embryo_120 * sc_skin;
  V_others    = V_embryo_120 * sc_others + V_embryo_0;
  V_muscle    = V_embryo_120 * sc_muscle;

  # Volume of the embryo without yolk (L)
  V_embryo = V_embryo_120 * (sc_liver + sc_gut + sc_skeleton + sc_eyes +
                             sc_brain + sc_heart + sc_skin + sc_others +
                             sc_muscle) + V_embryo_0;

  # Volume of yolk (L), decreases approximately exponentially with time
  V_yolk = V_yolk_0 * exp(-K_d_yolk * t);

  # Total volume of one embryo at start (L)
  V_embryo_total = V_yolk + V_embryo;

  # Total volume of medium + embryo, for one embryo (L)
  # only V_embryo_total is assumed to change with time
  V_content = V_medium_embryo + V_embryo_total;

  # Surface area of medium in contact with plastic per embryo,
  # with the embryo in the well (dm2)
  S_medium = 4 * V_content / D_well;

  # Volume of air (L) in head-space, per embryo
  V_air = V_well_embryo - V_content;

  # Apparent volume of distribution (L), per embryo
  V_app = V_medium_embryo / fu_diluted +
          P_aw * fui * V_air +
          P_pw * S_medium +
          P_mu_yolk * V_yolk +
          P_mu_liver * V_liver +
          P_mu_gut * V_gut +
          P_mu_muscle * V_muscle +
          P_mu_skeleton * V_skeleton +
          P_mu_eye * V_eyes +
          P_mu_brain * V_brain +

```

```

P_mu_heart      * V_heart      +
P_mu_skin       * V_skin       +
P_mu_others     * V_others;

# Sum of parent chemical's quantity in medium, air, plastic and embryo, per embryo
Q_parent = Q_labile + Q_fixed;

# Concentration unbound in medium (M) is (Q_labile - Q_air) over apparent V_medium
C_medium_u = Q_labile / ((V_medium_embryo / fu_diluted) + (P_aw * fui * V_air));

# Apparent volume of the embryo with yolk
V_app_e = V_app - V_medium_embryo / fu_diluted - P_aw * fui * V_air - P_pw *
S_medium;

# Quantity on plastic (mol),
Q_plastic = Q_fixed / (1 + fu_diluted * V_app_e / (P_pw * S_medium));

# Quantity and concentration in liver (M), null before organogenesis
tmp = P_mu_liver * V_liver;
Q_liver = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_liver = (V_liver > 0 ? Q_liver / V_liver : 0);

# Labile fraction (in medium and air) of Q_parent
f_labile = (V_medium_embryo / fu_diluted + P_aw * fui * V_air) / V_app;

# Linear metabolism in embryo, per embryo
N_cells = V_liver / V_cell_liver; # null before liver organogenesis
dt(Q_met) = (Michaelis > 0.5 ?
             N_cells * C_liver * Vmax / (Km + C_liver) : # MM
             N_cells * C_liver * K_met);                 # linear

dt(Q_labile) = k_diff * (Q_fixed - (1 / f_labile - 1) * Q_labile);
dt(Q_fixed) = -dt(Q_labile) - dt(Q_met);

}

CalcOutputs {

# Air, labile, depends on C_medium_u, computed in Dynamics
C_air      = P_aw * C_medium_u * fui; # concentration (M)
Q_air      = C_air * V_air;           # quantity (mol)

# Total in medium, labile, depends on C_medium_u, computed in Dynamics
C_medium   = C_medium_u / fu_diluted; # concentration (M)
Q_medium   = C_medium * V_medium_embryo;

# Other quantities part of Q_fixed in mol (per embryo)

Q_parent   = Q_labile + Q_fixed;

# Q_plastic is computed in Dynamics
C_plastic  = Q_plastic / S_medium; # (mol/dm^2)

# Quantity and concentration in organs (M), null before organogenesis
# the needed parameters and variables have been computed in Dynamics

# yolk
tmp        = P_mu_yolk * V_yolk;
Q_yolk     = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_yolk     = (V_yolk > 0 ? Q_yolk / V_yolk : 0);

# Q_liver and C_liver are computed in dynamics

# gut
tmp        = P_mu_gut * V_gut;
Q_gut      = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_gut      = (V_gut > 0 ? Q_gut / V_gut : 0);

```

```

# muscle
tmp      = P_mu_muscle * V_muscle;
Q_muscle = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_muscle = (V_muscle > 0 ? Q_muscle / V_muscle : 0);

# skeleton
tmp      = P_mu_skeleton * V_skeleton;
Q_skeleton = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_skeleton = (V_skeleton > 0 ? Q_skeleton / V_skeleton : 0);

# eyes
tmp      = P_mu_eye * V_eyes;
Q_eyes   = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_eyes   = (V_eyes > 0 ? Q_eyes / V_eyes : 0);

# brain
tmp      = P_mu_brain * V_brain;
Q_brain  = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_brain  = (V_brain > 0 ? Q_brain / V_brain : 0);

# heart
tmp      = P_mu_heart * V_heart;
Q_heart  = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_heart  = (V_heart > 0 ? Q_heart / V_heart : 0);

# skin
tmp      = P_mu_skin * V_skin;
Q_skin   = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_skin   = (V_skin > 0 ? Q_skin / V_skin : 0);

# others
tmp      = P_mu_others * V_others;
Q_others = (Q_fixed - Q_plastic) / (1 + (V_app_e - tmp) / tmp);
C_others = (V_others > 0 ? Q_others / V_others : 0);

Q_embryo = Q_liver + Q_gut + Q_muscle + Q_skeleton + Q_eyes + Q_brain + Q_heart +
Q_skin + Q_others;

C_embryo = Q_embryo / V_embryo;

Q_embryo_total = Q_embryo + Q_yolk;

C_embryo_total = Q_embryo_total / V_embryo_total;

# Concentration in lysosomes and mitochondria are anchored to C_others
C_lyso      = P_mu_ly * C_others / P_mu_others;
C_mito      = P_mu_mi * C_others / P_mu_others;

Q_lyso      = C_lyso * V_embryo * f_lyso;
Q_mito      = C_mito * V_embryo * f_mito;

# Safeguards against negative of null values
C_medium    = (C_medium > 0 ? C_medium : 1E-30);
C_embryo_total = (C_embryo_total > 0 ? C_embryo_total : 1E-30);

Q_check = Q_medium + Q_embryo + Q_yolk + Q_air + Q_plastic + Q_met;
# an alternative would be Q_labile + Q_fixed...

# Length of the embryo (dm), calibrated with data from Kimmel 1995
# This could be improved and made temperature dependent
A = 2.6039428;
B = 4.3973752;
C = 26.9499832;
D = 0.7551494;
L_embryo = (A * pow(t/60, B) / (pow(C, B) + pow(t/60, B)) + D) / 100;
} End.

```

Annexe 3 : Supplementary Material (Article modèle PK-PD multi-états)

Section 1 – Experimental data

Table S1. Codes of effects recorded in the Zebrafish Embryo Acute Toxicity (FET) test.

Code string	Meaning
*	Outlier (for technical reasons)
N	Normal development
HA	Hatched
C	Coagulation
TA	Tail attached
GA	Gastrulation arrest
S	Somite malformation
H-	Reduced heart beat
H=	Severely reduced heart beat
H	Missing heart beat
B	Blood pooling
BD	Delayed body development
BL(P)-	Reduced blood flow with pooling at yolk
BL(P)=	Severely reduced blood flow with pooling at yolk
BL(P)	Missing blood flow with pooling at yolk
BL-	Reduced blood flow
BL=	Severely reduced blood flow
BL	Missing blood flow
BR	Visible brain
CH	Chorda unclear
DY	Dark yolk
E	Edema
PE	Pericardial edema
YE	Yolk edema
EB	Eyes larger than usual
ED-	Reduced eye development
ED=	Severely reduced eye development
ED	Missing eye development
EP-	Reduced eye pigmentation

Code string	Meaning
EP=	Severely reduced eye pigmentation
EP	Missing eye pigmentation
FM	Missing face
FU	Fin underdeveloped
GA	Gastrulation arrest
HD	Head deformed
HZ	Head tremor/twitching
JD	Craniofacial deformation
L	Lordosis
KY	Kyphosis
SC	Scoliosis
M	Motionless
MD	Muscle development/ boundaries unclear
NH	Not hatched
O	Otolith development impacted
P-	Reduced pigmentation
P=	Severely reduced pigmentation
P	Missing pigmentation
PF	Pectoral fins not moving
PFA	Pectoral fins attached
PFD	Pectoral fins deformed
SL	Side lying
SM	Spontaneous movement missing
TD	Tailfin decomposing
TF	Tailfin development impacted
YD	Yolk deformation
YR	Reduced yolk resorption
Z	Tremor/twitching

Table S2. Water pH after addition of valproic acid (VPA) at various test concentrations. At the highest concentration, all embryos died within a few hours. Prior to the addition of the test compound, water was adjusted at pH 7.75 (± 0.02).

VPA CONCENTRATION (μM)	PH
800	5.11
400	6.23
200	6.83
100	7.11
50	7.32
25	7.41
12.5	7.55
6.25	7.64

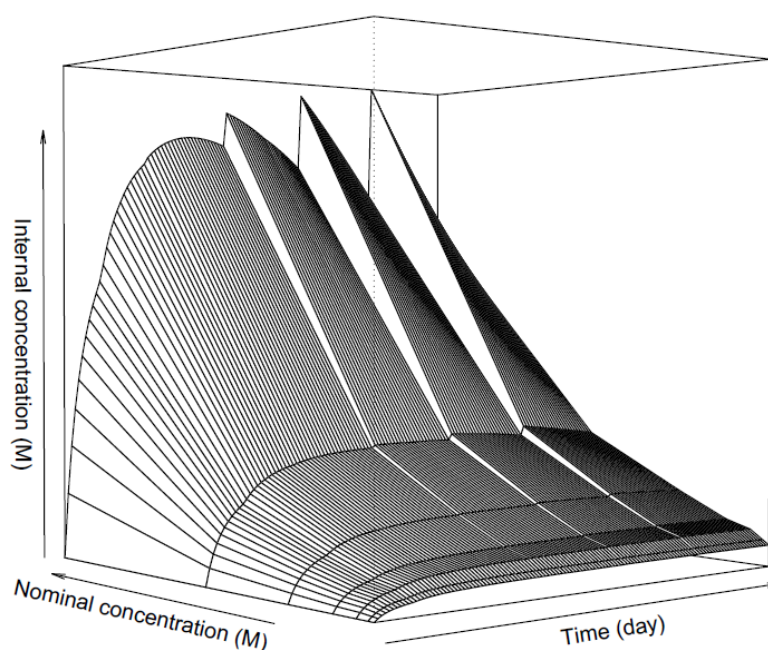


Figure S1. Internal concentration (ranging from 0 to 1.6×10^{-2} M) as a function of the nominal concentration (ranging from 0 to 4×10^{-4} M) and the time (from 0 to 5 days).

Section 2 – Hatching rate modeling

We tested five sub-models linking $q_{N,H}(t)$ and time to fit the data reasonably well. We only used the data of the control group and a three-state model, considering only normal, hatched, and malformed or coagulated embryos, because in the control group the latter state is rarely observed, and hatching occurs most of the time. In that group we therefore have quite clean observations of the hatching process alone. The model structure is illustrated in Figure S2, below.

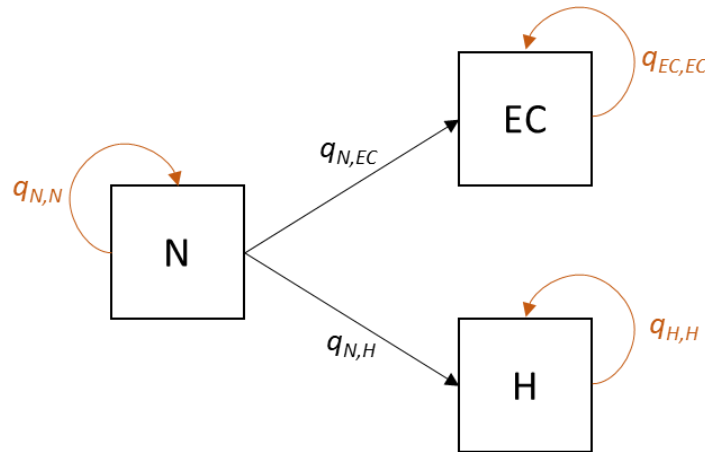


Figure S2. Schematic representation of the three-state model for zebrafish embryo hatching. States are: “normal” (pre-hatching, N), “hatched” (H), and “all malformed or coagulated embryos” (EC).

The five sub-models describing $q_{N,H}(t)$ time-dependence were:

- A constant rate model (*a priori* probably wrong), with two parameters to estimate ($q_{N,H}$ and $q_{N,EC}$)
- A piecewise constant model with a fixed hatching time at 48 hpf, with two parameters to estimate ($q_{N,H}$ after 48 hpf and $q_{N,EC}$; $q_{N,H}$ before 48 hpf was set to zero)
- A piecewise constant model with variable hatching time t_h , making three parameters to estimate (t_h , $q_{N,H}$ after t_h , and $q_{N,EC}$; $q_{N,H}$ before t_h being null)
- A Hill’s model such that the transition rate $q_{H,N}(t)$ is given by:

$$q_{N,H}(t) = \frac{q_{N,H_{max}} \times t^n}{t_{50}^n + t^n} \quad (S1)$$

where $q_{N,H_{max}}$ is the maximum transition intensity from normal to hatched, t_{50} is the time at which the transition rate is equal to 50% of its maximum value, and n is the Hill’s exponent. Those three parameters and $q_{N,EC}$ were to be estimated.

- A Gaussian-form model, reaching a maximum and then decaying, so that $q_{N,H}(t)$ is given by:

$$q_{N,H}(t) = \frac{A}{s} \times e^{\left(\frac{t-t_h}{s}\right)^2} \quad (S2)$$

where t_h is the mean hatching time, s its spread, and A/s the maximum rate, so four parameters were estimated: t_h , s , A , and $q_{N,EC}$.

To choose between those models based on their fit to the data, we used the Akaike information criterion (AIC). It uses the maximum likelihood, $\widehat{L}_{l,j}$, as a measure of goodness-of-fit, penalized by the number k of estimated parameters (Evans, 2019):

$$AIC = -2 \times \ln(\widehat{L}_{l,j}) + 2 \times k \quad (S3)$$

The model with the lowest AIC should be preferred.

The posterior distribution of parameters obtained by MCMC calibration with the embryo transition count data, are summarized for each model in Table S4. The AIC values were 570 for the constant model, 287 for the piecewise constant model with fixed hatching time, 269 for the piecewise constant model with estimated hatching time, 269 for the Hill's model, and 276 for the Gaussian model. The piecewise constant model with estimated hatching time was therefore the best model according to the AIC and was selected (they all give practically the same fit, but it is essentially the most parsimonious).

In the five-state model (Figure 3 of the main text), hatching can happen to normal embryos (state N) and malformed embryos (state E). The piecewise constant sub-model with estimated hatching time was therefore used for both $q_{N,H}$ and $q_{E,EH}$ hatching intensities in the final model.

Table S3. Posterior mean, standard deviation (SD), 95% credibility intervals (IC95%) and maximum posterior (MP) value of the parameters for the five hatching rate sub-model in control embryos.

Parameter	Mean $\pm$ SD	IC 95%	MP	$\hat{R}$
Constant transition rate				
$q_{N,H}$	0.35 ± 0.033	[0.29 ; 0.41]	0.34	1
$q_{N,EC}$	0.037 ± 0.020	[0.019 ; 0.061]	0.033	1
Hatching rate piecewise constant with fixed hatching time				
$q_{N,H}$	1.3 ± 0.13	[1.0 ; 1.6]	1.3	1
$q_{N,EC}$	0.038 ± 0.024	[0.019 ; 0.061]	0.034	1
Hatching rate piecewise constant with estimated hatching time				
$q_{N,H}$	3.3 ± 0.86	[2.0 ; 5.3]	3.2	1
$q_{N,EC}$	0.035 ± 0.012	[0.018 ; 0.058]	0.032	1
t_h (day)	2.7 ± 0.090	[2.5 ; 2.8]	2.7	1
Hatching rate following a Hill function				
$q_{N,Hmax}$	7.2 ± 2.7	[3.0 ; 13]	9.0	1
$q_{N,EC}$	0.036 ± 0.010	[0.019 ; 0.059]	0.034	1
n	6.2 ± 1.6	[3.6 ; 9.5]	4.9	1
x	3.4 ± 0.35	[2.7 ; 4.0]	3.8	1
Hatching rate following a Gaussian function				
$q_{N,EC}$	0.036 ± 0.010	[0.019 ; 0.059]	0.034	1
s	0.89 ± 0.14	[0.64 ; 1.2]	0.82	1
A	5.8 ± 1.9	[2.9 ; 9.6]	4.4	1
t_h (day)	3.8 ± 0.23	[3.5 ; 4.3]	3.7	1

Figure S3 shows the transition number of zebrafish embryos from “N” to “N”, “EC”, and “H” states, for the five hatching sub-models, as a function of the time (day). Except for the constant intensity model, the data are reasonably well fitted.

Figure S4 shows the models’ predictions of time evolution of embryo counts in each of the three state.

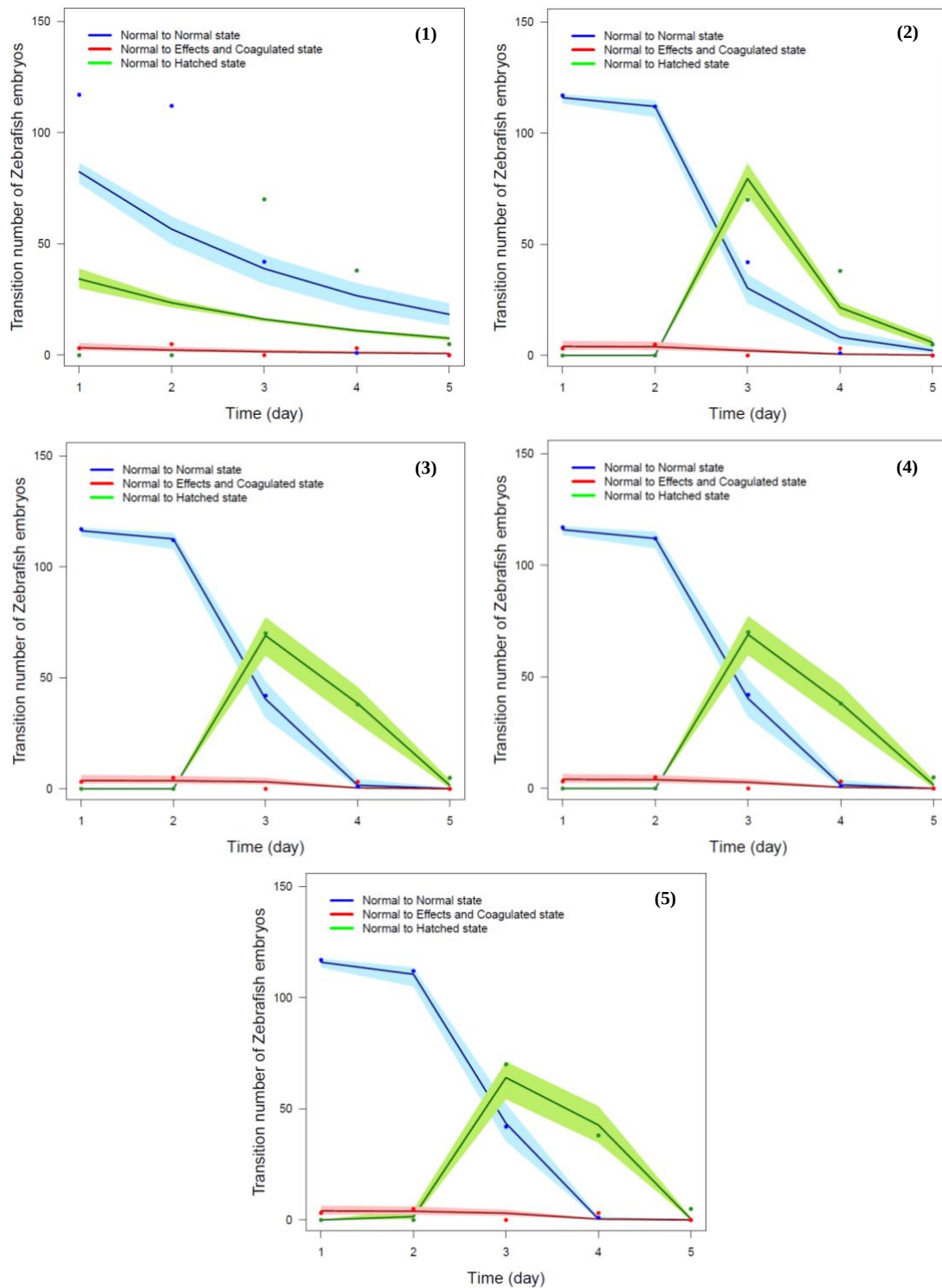


Figure S3. Number of transitions between states N, EC and H, as a function of time. Hatching was modeled by (1) a constant intensity sub-model, (2) a piecewise constant model with fixed 48 hpf hatching time, (3) a piecewise constant model with estimated hatching time, (4) a Hill's model, and (5) a Gaussian model. The dots are data, and the lines predictions. The blue, red and green areas are the 95% confidence interval of transition number from N to N, from N to EC, and from N to H.

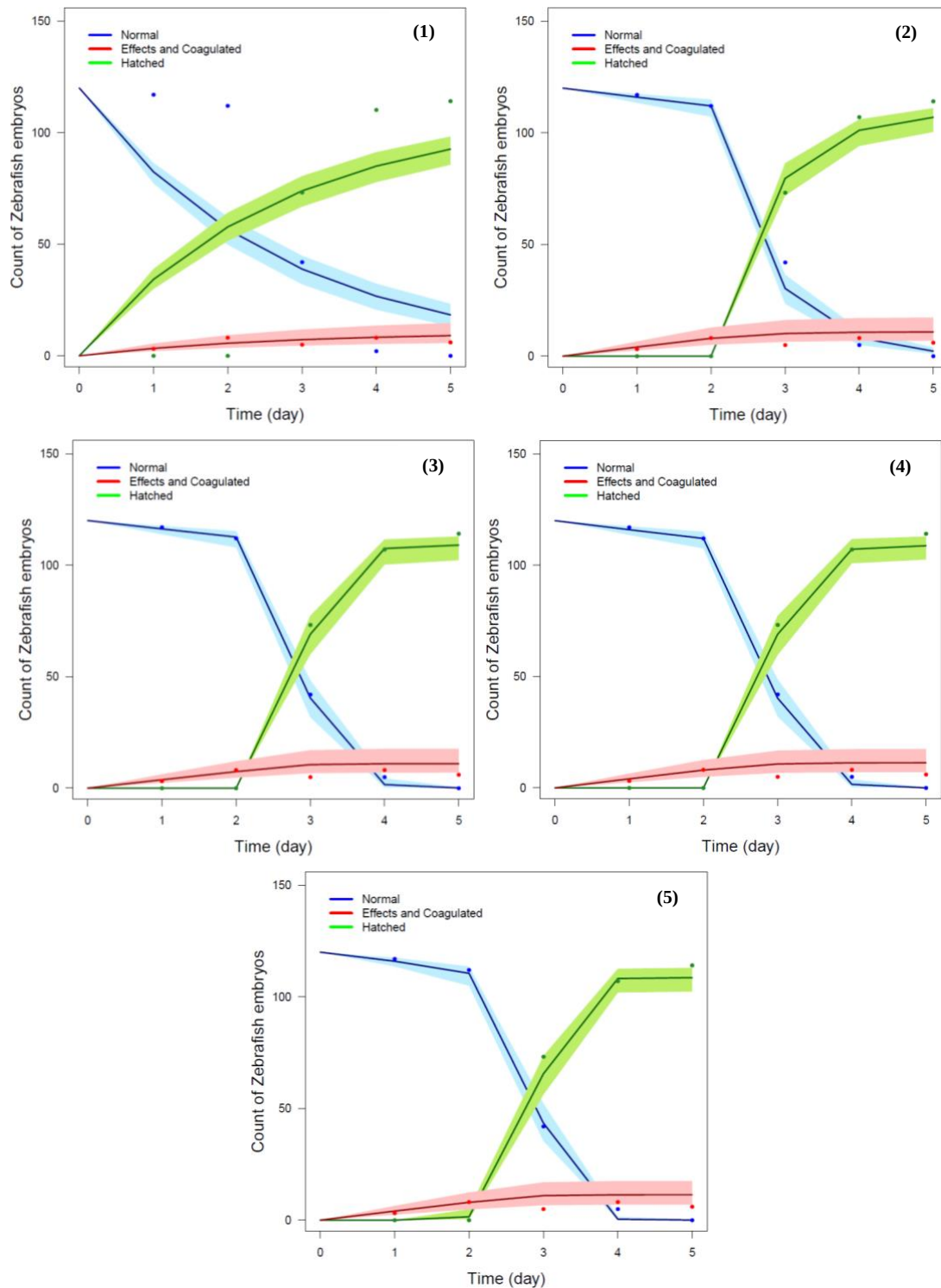


Figure S4. Fish count in states N, EC and H, as a function of time. Hatching was modeled by (1) a constant intensity sub-model, (2) a piecewise constant model with fixed 48 hpf hatching time, (3) a piecewise constant model with estimated hatching time, (4) a Hill's model, and (5) a Gaussian model. The dots are data, and the lines predictions. The blue, red and green areas are the 95% confidence interval of Fish count in states N, EC and H.

Figure S5 shows the hatching intensity $q_{N,H}(t)$ according to the various models, except for the constant rate model that poorly fitted the data. The profiles are very different according to the model used, but it note that all models pass by the point (2.7, 3.3), which seems pivotal for a reasonable description of the data. There is a large uncertainty about the later part of the curve for the two most complex models, because after day 3 there are very few non-hatched fish left to inform the exact shape of the curve.

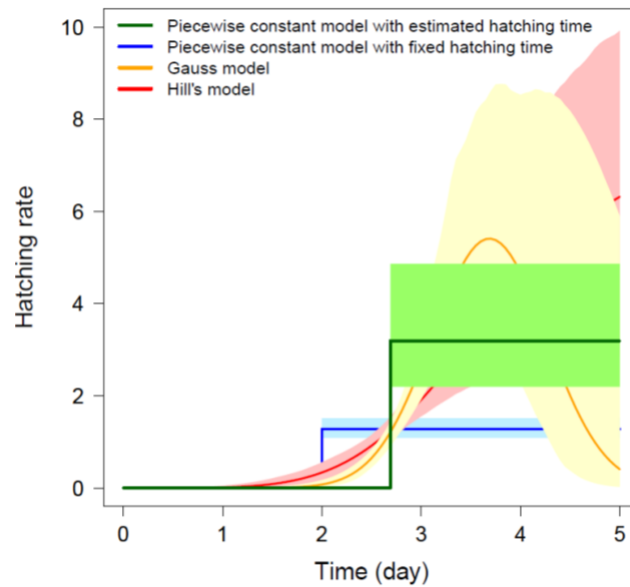


Figure S5. Shape of the time-dependent hatching intensity with the piecewise constant with estimated hatching time (green), the piecewise constant with fixed hatching time (blue), the Gaussian (orange), and the Hill's (red) three-state model. The colored areas are the 95% credibility interval of hatching intensity estimates. The thick lines are the maximum posterior hatching rate functions.

Section 3 – Coupled PBPK-multistate model results

Table S4. Observed and predicted value of transition number for each day and concentration. The most unexpected observations (according to the model) are in bold font.

Concentration (μM)	Observations					Predictions				
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 1	Day 2	Day 3	Day 4	Day 5
<i>Transition $N \rightarrow N$</i>										
0	117	112	42	1	0	116.6	113.2	45	1.7	0.07
6.25	60	53	24	0	0	57.8	55.5	22.3	0.79	0.38
12.5	60	58	26	0	0	57.7	55.2	21	0.77	0.03
25	59	55	24	0	0	57.4	54.5	25.1	0.89	0.04
50	59	54	31	1	1	57.2	53.8	26.8	0.95	0.043
100	57	50	25	0	0	56.2	51.3	24.1	0.82	0.038
200	50	38	19	0	0	52.5	41.6	25.6	0.76	0.034
400	0	0	0	0	0	3.5×10^{-2}	6.4×10^{-6}	1.4×10^{-6}	9.2×10^{-7}	9.1×10^{-6}
<i>Transition $N \rightarrow C$</i>										
0	0	0	0	0	0	1.4×10^{-2}	1.3×10^{-2}	1.2×10^{-2}	1.6×10^{-3}	6.8×10^{-5}
6.25	0	0	0	0	0	0.03	0.04	0.038	0.0053	0.00026
12.5	0	0	0	0	0	0.035	0.048	0.044	0.0054	0.00021
25	1	0	0	0	0	0.045	0.063	0.061	0.0095	0.00043
50	1	0	0	0	0	0.054	0.077	0.075	0.012	0.00057
100	0	0	0	0	0	0.086	0.13	0.12	0.019	0.0009
200	0	0	0	0	0	0.18	0.29	0.26	0.055	0.0025
400	0	0	0	0	0	2.7	0.0031	0.00093	0.00056	0.00049
<i>Transition $N \rightarrow E$</i>										
0	3	5	0	0	0	3.5	3.4	2.4	0.19	0.0082
6.25	0	7	0	0	0	2.2	2.4	1.9	0.19	0.0091
12.5	0	2	1	0	0	2.4	2.6	1.2	0.058	0.0022
25	0	4	0	0	0	2.6	2.9	2.3	0.22	0.0099
50	0	5	4	0	0	2.9	3.3	2.7	0.27	0.012
100	3	7	3	1	0	3.8	4.8	3.8	0.35	0.016
200	10	12	14	4	0	7.4	10.6	8.7	0.90	0.041
400	60	0	0	0	0	57.4	0.032	0.0092	0.0053	0.0048
<i>Transition $N \rightarrow H$</i>										
0	0	0	70	38	5	0	0	66.9	44	1.9
6.25	0	0	29	25	0	0	0	31.9	21.9	1
12.5	0	0	30	25	0	0	0	32.7	20.3	0.78
25	0	0	31	24	0	0	0	27.3	23.6	1
50	0	0	18	22	0	0	0	24.4	23	1
100	0	0	15	14	0	0	0	18.8	6.9	0.31
200	0	0	8	0	0	0	0	4.1	0.74	0.035
400	0	0	0	0	0	0	0	2.9×10^{-6}	1.2×10^{-5}	1.3×10^{-5}
<i>Transition $N \rightarrow EH$</i>										
0	0	0	0	3	0	0	0	0.55	0.16	0.0067
6.25	0	0	0	0	0	0	0	0.21	0.13	0.0066
12.5	0	0	1	1	0	0	0	1.1	0.40	0.016
25	0	0	0	2	0	0	0	0.6	1.3	0.059
50	0	0	1	8	0	0	0	0.9	3.5	0.16
100	0	0	7	10	0	0	0	5.6	17.1	0.79
200	0	0	1	15	0	0	0	4.3	25.2	1.12
400	0	0	0	0	0	0	0	0.00028	0.0011	0.00095

Concentration (μM)		Observations					Predictions				
Transition $E \rightarrow C$											
0	0	0	0	0	0	0	0.00021	0.0075	0.011	0.0070	0.0019
6.25	0	0	0	0	0	0	0.00041	0.012	0.021	0.019	0.0085
12.5	0	0	0	0	0	0	0.00046	0.015	0.024	0.0066	0.00052
25	0	0	0	0	0	0	0.00055	0.02	0.036	0.028	0.01
50	0	0	0	0	0	0	0.00062	0.026	0.048	0.039	0.014
100	0	0	0	0	2	0	0.00089	0.051	0.10	0.081	0.028
200	0	0	0	0	0	0	0.0015	0.18	0.42	0.35	0.11
400	0	1	1	18	0	0	0.006	5.1	4.7	3.9	2.9
Transition $E \rightarrow N$											
0	0	0	0	0	0	0	0.045	1.6	1.1	0.31	0.085
6.25	0	0	1	0	0	0	0.041	0.90	0.67	0.3	0.13
12.5	0	0	0	0	0	0	0.041	0.92	0.62	0.054	0.0042
25	0	0	2	0	0	0	0.039	0.96	0.84	0.27	0.092
50	0	0	0	0	0	0	0.038	1	0.95	0.3	0.11
100	0	0	0	0	0	0	0.034	1.1	1.1	0.34	0.11
200	0	4	0	0	0	0	0.025	1.3	2.1	0.48	0.16
400	0	0	0	0	0	0	5.6×10^{-5}	0.01	0.0069	0.0063	0.054
Transition $E \rightarrow E$											
0	0	3	4	2	2	0	0.055	1.9	2.3	1.1	0.31
6.25	0	0	4	0	0	0	0.058	1.4	2	1.6	0.72
12.5	0	0	0	0	0	0	0.059	1.5	1.2	0.14	0.011
25	0	0	1	1	1	0	0.060	1.7	2.5	1.5	0.51
50	0	0	1	1	0	0	0.062	1.9	2.9	1.8	0.64
100	0	3	4	3	0	0	0.065	2.7	4.5	2.5	0.84
200	0	6	16	2	3	0	0.073	6	12	6	2
400	0	59	57	34	29	0	0.094	52.4	46.3	36.4	28.6
Transition $E \rightarrow H$											
0	0	0	3	1	0	0	0	0	1.3	1.7	0.47
6.25	0	0	2	4	0	0	0	0	0.8	1.2	0.55
12.5	0	0	2	1	0	0	0	0	0.90	0.84	0.066
25	0	0	0	0	0	0	0	0	0.81	1.4	0.49
50	0	0	3	2	0	0	0	0	0.78	1.51	0.54
100	0	0	2	0	0	0	0	0	0.79	1	0.35
200	0	0	0	0	0	0	0	0	0.34	0.40	0.14
400	0	0	0	0	0	0	0	0	0.015	0.069	0.065
Transition $E \rightarrow EH$											
0	0	0	1	1	0	0	0	0	0.61	1.6	0.44
6.25	0	0	0	0	0	0	0	0	0.23	0.81	0.37
12.5	0	0	0	0	0	0	0	0	1.2	1.4	0.11
25	0	0	1	0	0	0	0	0	0.45	1.7	0.58
50	0	0	1	2	1	0	0	0	0.47	2.1	0.74
100	0	0	4	4	2	0	0	0	1	4.4	1.5
200	0	0	2	28	3	0	0	0	1.3	13.9	4.5
400	0	0	1	5	5	0	0	0	1.4	6	4.7
Transition $H \rightarrow H$											
0	0	0	0	72	106	0.1	0.14	0.16	68.4	114.5	
6.25	0	0	0	31	60	0.1	0.14	0.16	32.7	55.7	
12.5	0	0	0	31	57	0.1	0.13	0.16	33.4	54.7	
25	0	0	0	31	55	0.099	0.13	0.14	26.6	48.7	
50	0	0	0	16	34	0.097	0.12	0.12	21.3	38.5	
100	0	0	0	3	2	0.071	0.031	0.012	4	2.6	
200	0	0	0	0	0	0.0032	0.00014	0.00010	0.098	0.032	
400	0	0	0	0	0	1.1×10^{-3}	2.5×10^{-5}	2.4×10^{-5}	1.9×10^{-4}	1.1×10^{-3}	

Concentration (μM)	Observations					Predictions					
<i>Transition $H \rightarrow EH$</i>											
0	0	0	0	1	1	0	0	0	0	0	0
6.25	0	0	0	0	0	6.5×10^{-5}	3.4×10^{-4}	5.5×10^{-4}	0.14	0.25	
12.5	0	0	0	1	0	0.00018	0.00096	0.0015	0.4	0.7	
25	0	0	0	0	0	0.00096	0.0049	0.0075	1.7	3.2	
50	0	0	0	5	6	0.0028	0.013	0.019	4	7.7	
100	0	0	0	14	15	0.029	0.066	0.038	15.6	10.3	
200	0	0	0	8	0	0.097	0.006	0.0046	4.4	1.3	
400	0	0	0	0	0	0.099	0.0023	0.0021	0.016	0.085	
<i>Transition $EH \rightarrow H$</i>											
0	0	0	0	0	3	0.036	0.023	0.015	0.43	0.9	
6.25	0	0	0	0	0	0.036	0.023	0.015	0.17	0.49	
12.5	0	0	0	0	3	0.035	0.023	0.015	0.83	1.3	
25	0	0	0	0	2	0.035	0.023	0.016	0.37	1.8	
50	0	0	0	0	9	0.034	0.022	0.019	0.45	3.3	
100	0	0	0	0	2	0.027	0.018	0.023	0.96	6.1	
200	0	0	0	0	0	0.003	0.0045	0.0043	0.13	1.1	
400	0	0	0	0	0	0.0011	0.0021	0.0022	0.018	0.095	
<i>Transition $EH \rightarrow EH$</i>											
0	0	0	0	1	3	0.064	0.041	0.026	0.76	1.6	
6.25	0	0	0	0	0	0.064	0.042	0.027	0.3	0.9	
12.5	0	0	0	1	0	0.065	0.042	0.028	1.5	2.4	
25	0	0	0	1	1	0.065	0.043	0.032	0.72	3.5	
50	0	0	0	2	8	0.66	0.046	0.04	0.98	7.2	
100	0	0	0	11	37	0.073	0.084	0.13	5.8	36.8	
200	0	0	0	3	54	0.097	0.19	0.19	5.7	48	
400	0	0	0	1	6	0.099	0.20	0.20	1.6	7.5	

Table S5. Mean, standard deviation (SD), 95% confidence intervals (IC95%) and maximum posterior (MP) value for concentration-group-specific estimates of $q_{E,EH}$, $q_{N,H}$, and t_h .

Parameter (0 μ M)	Mean $\pm$ SD	IC 95%	MP	$\hat{R}$
Concentration 0 μ M				
$q_{E,EH}$	0.79 ± 0.32	[0.32 ; 1.5]	0.83	1
$q_{N,H}$	3.8 ± 0.62	[2.8 ; 5.3]	3.3	1
t_h	2.7 ± 0.049	[2.6 ; 2.8]	2.7	1
Concentration 6.25 μ M				
$q_{E,EH}$	0.82 ± 0.60	[0.15 ; 2.4]	0.40	1
$q_{N,H}$	4.2 ± 1.1	[2.9 ; 7.1]	3.4	1
t_h	2.8 ± 0.049	[2.7 ; 2.9]	2.7	1
Concentration 12.5 μ M				
$q_{E,EH}$	1.6 ± 1.4	[0.29 ; 5.0]	2.4	1
$q_{N,H}$	4.1 ± 1.0	[2.8 ; 6.5]	3.3	1
t_h	2.8 ± 0.050	[2.7 ; 2.9]	2.7	1
Concentration 25 μ M				
$q_{E,EH}$	0.71 ± 0.39	[0.19 ; 1.7]	0.74	1
$q_{N,H}$	4.0 ± 0.93	[2.8 ; 6.4]	3.3	1
t_h	2.8 ± 0.050	[2.7 ; 2.9]	2.8	1
Concentration 50 μ M				
$q_{E,EH}$	1.4 ± 0.65	[0.48 ; 3.0]	0.74	1
$q_{N,H}$	3.4 ± 0.56	[2.3 ; 4.6]	3.3	1
t_h	2.8 ± 0.039	[2.7 ; 2.9]	2.8	1
Concentration 100 μ M				
$q_{E,EH}$	1.1 ± 0.38	[0.47 ; 2.0]	0.86	1
$q_{N,H}$	3.4 ± 0.59	[2.3 ; 4.7]	3.4	1
t_h	2.8 ± 0.045	[2.7 ; 2.9]	2.8	1
Concentration 200 μ M				
$q_{E,EH}$	1.4 ± 0.30	[0.87 ; 2.1]	1.0	1
$q_{N,H}$	3.8 ± 0.86	[2.6 ; 5.9]	3.4	1
t_h	2.9 ± 0.025	[2.9 ; 3.0]	2.9	1
Concentration 400 μ M				
$q_{E,EH}$	0.17 ± 0.045	[0.094 ; 0.27]	0.15	1
$q_{N,H}$	3.9 ± 1.2	[2.2 ; 7.1]	3.3	1
t_h	2.8 ± 0.077	[2.7 ; 3.0]	2.8	1

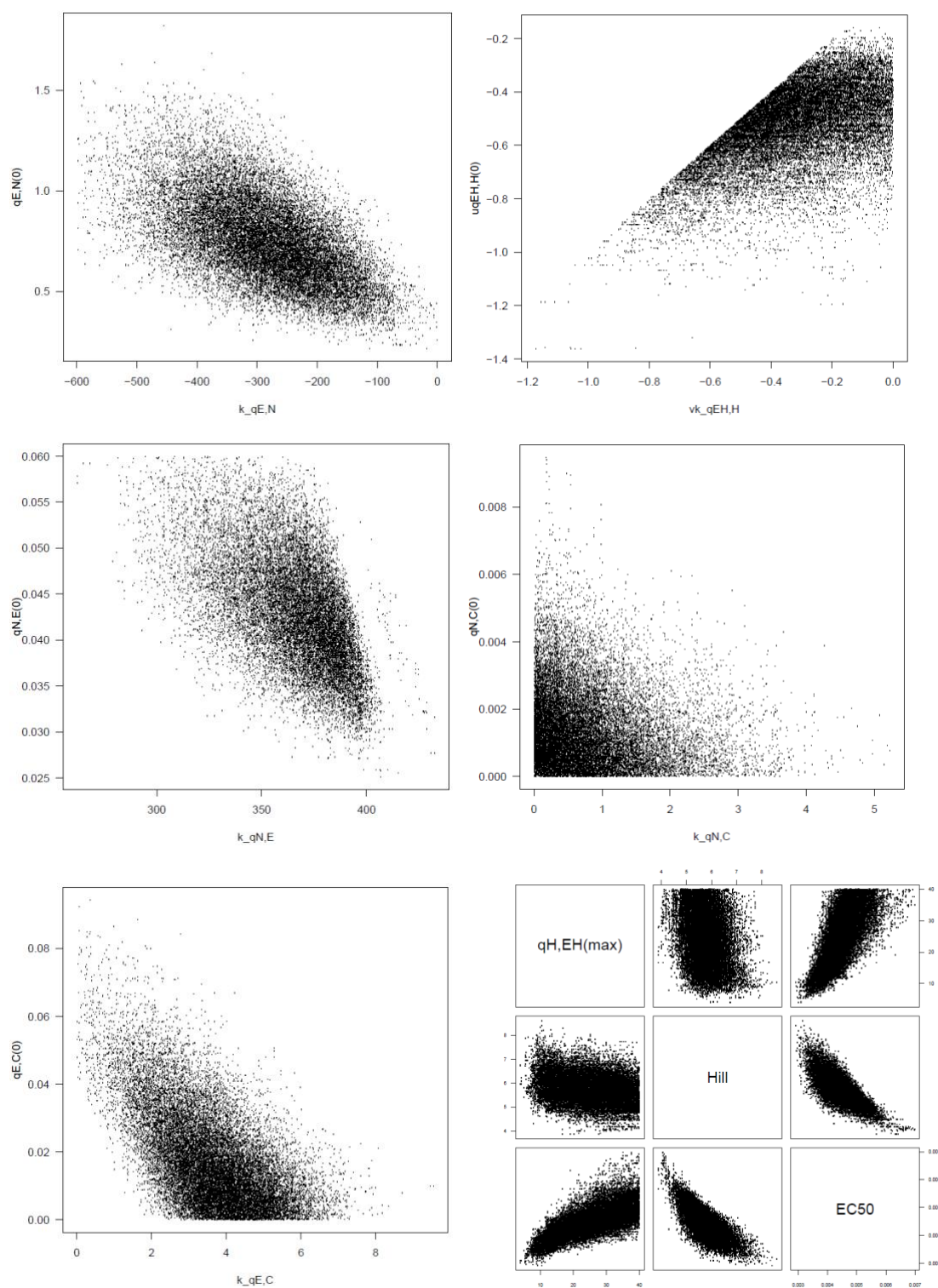


Figure S6. Correlation of the parameter estimates.

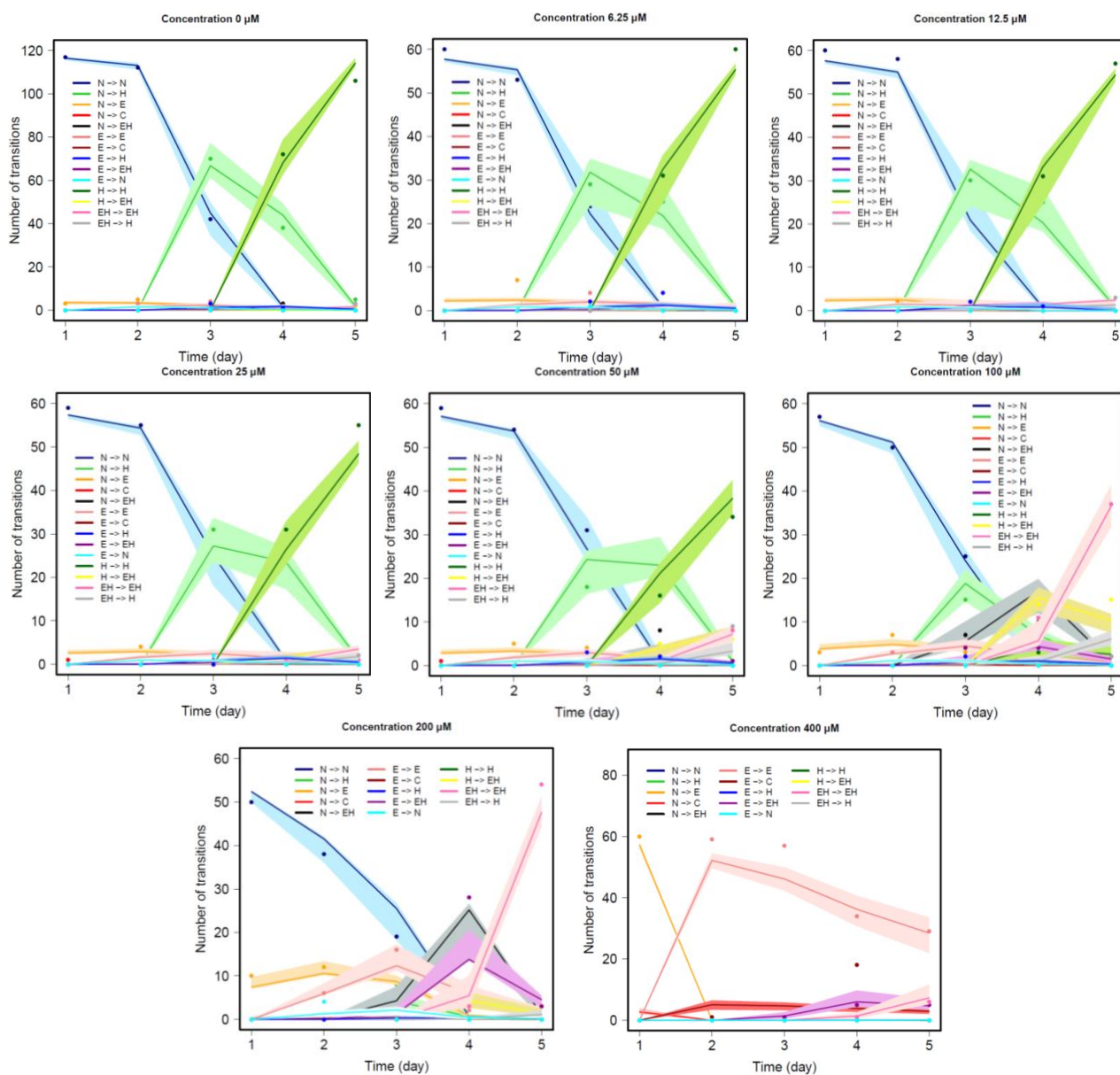


Figure S7. Observed (dots) and predicted (lines) number of transitions between states normal (N), hatched (H), malformed (E), hatched with effects (EH) and coagulated (C), as a function of time for the exposure concentrations 0 and 400 μM . Hatching rate is modeled by a piecewise constant intensity with estimated hatching time. The colored bands give the predictions' 95% confidence intervals.

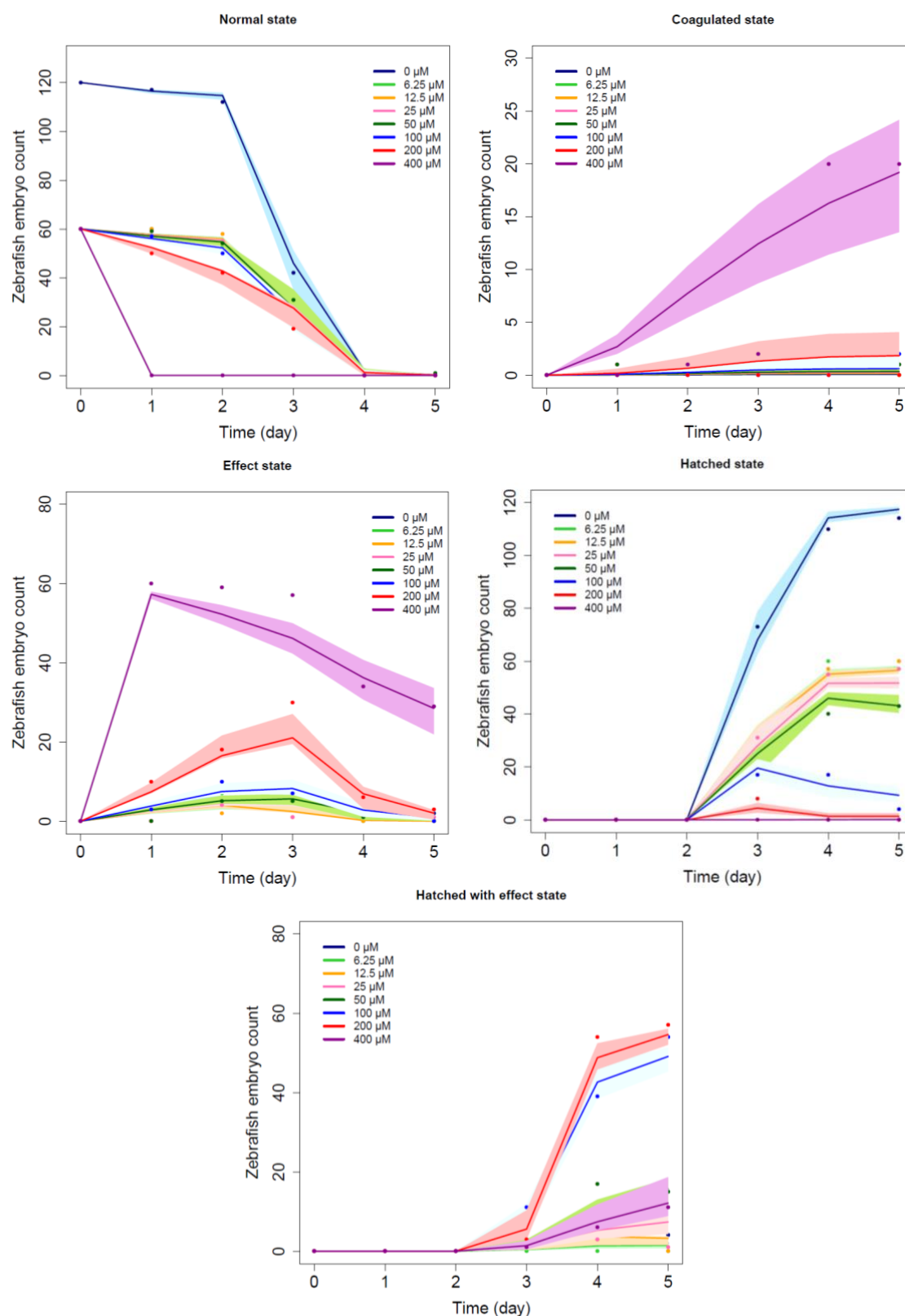


Figure S8. Counts of embryos in the five states, as a function of time and VPA concentrations. The dots are the data, the lines are model predictions. The colored bands give the predictions' 95% confidence intervals.

Section 4 – Calculation of reference toxicity values

The transition intensities do not depend directly on time, if we take them as functions of internal concentration (Eqs. 28 to 33). In that case, given the posterior distribution of the parameters of those equations it is easy, by mathematical inversion, to compute the distributions of the reference toxicity values of interest. According to Equations 28 to 33 of the main text, the EC_{10} values for $q_{E,N}$, $q_{EH,H}$, $q_{N,E}$, $q_{N,C}$, $q_{E,C}$, and $q_{H,EH}$ – respectively, $EC_{10}(q_{E,N})$, $EC_{10}(q_{EH,H})$, $EC_{10}(q_{N,E})$, $EC_{10}(q_{N,C})$, $EC_{10}(q_{E,C})$, and $EC_{10}(q_{H,EH})$ – are given respectively by:

$$q_{i,j}(EC_{10}) = q_{i,j}(0) + 0.1 \times (q_{i,j}(Cmax) - q_{i,j}(0)) \quad (S4)$$

where $q_{i,j}(Cmax)$ is the transition rate corresponding to the maximal concentration value; and:

$$EC_{10}(q_{E,N}) = \frac{\log(q_{E,N}(EC_{10})) - \log(q_{E,N}(0))}{k_{q_{E,N}}} \quad (S5)$$

$$EC_{10}(q_{EH,H}) = \frac{q_{EH,H}(EC_{10}) - q_{EH,H}(0)}{k_{q_{EH,H}}} \quad (S6)$$

$$EC_{10}(q_{N,E}) = \frac{\log(q_{N,E}(EC_{10})) - \log(q_{N,E}(0))}{k_{q_{N,E}}} \quad (S7)$$

$$EC_{10}(q_{N,C}) = \frac{q_{N,C}(EC_{10}) - q_{N,C}(0)}{k_{q_{N,C}}} \quad (S8)$$

$$EC_{10}(q_{E,C}) = \frac{q_{E,C}(EC_{10}) - q_{E,C}(0)}{k_{q_{E,C}}} \quad (S9)$$

$$EC_{10}(q_{H,EH}) = \sqrt[n]{\frac{q_{H,EH}(EC_{10}) \times EC_{50}^n}{q_{H,EH_{max}} - q_{H,EH}(EC_{10})}} \quad (S10)$$

Table S6. Effect concentrations (EC), no effect concentration (NOEC) and lowest effect concentration (LOEC) of all effects caused by valproic acid in zebrafish (*Danio rerio*) embryos after 120 h of exposure. Those were derived only on the basis of data at time 120 hpf.

Effects	NOEC (μM)	LOEC (μM)	EC ₁₀ (μM)	EC ₅₀ (μM)
Coagulation	12.5	400	88	335
Pericardial edemata	25	50	38	90
Yolk edemata	12.5	25	82	<i>n.d.</i>
Yolk sac absorption reduced	12.5	25	20	194
Reduced heartbeat	200	400	209	<i>n.d.</i>
Blood congestion	100	200	119	<i>n.d.</i>
Scoliosis/lordosis	6.25	50	59	230
Craniofacial deformation	50	100	52	142
Small eyes	200	400	<i>n.d.</i>	<i>n.d.</i>
Jitter/tremor	50	100	82	196
Development retardation	200	400	<i>n.d.</i>	<i>n.d.</i>
Lack of hatch	200	400	300	<i>n.d.</i>

Those estimates were obtained with ToxRat Professional version 2.10.03 (ToxRat Solutions, Alsdorf, Germany) on the basis of the total number of affected embryos per test concentration at 120 hpf. For calculating EC values concentration 800 μM was excluded for all effects except for coagulation (death).

n.d.: Not determined (> highest test concentration)

Section 5 – Comparing R package *msm* and GNU MCSim results for CAV

We used the coronary allograft vasculopathy (CAV) multistate model and data described in Jackson and available in the *msm* R package (Jackson, C.H., 2011. Multi-State Models for Panel Data: The *msm* Package for R. J. Stat. Soft. 38. <https://doi.org/10.18637/jss.v038.i08>). Figure S9 shows the parameter samples obtained by MCMC sampling with *GNU MCSim*, superimposed with the maximum likelihood and confidence interval obtained by bootstrapping with *msm*. *GNU MCSim* took 20 minutes to obtain 2000 MCMC samples, and *msm* took 1 hour to obtain 1000 bootstrap samples. The convergence of the MCMC chains is very fast. GNU MCSim MCMCs and *msm* bootstrapping are in complete agreement. The normal (Hessian) approximation of the 95% confidence intervals of *msm* can be large over-estimates (see q24).

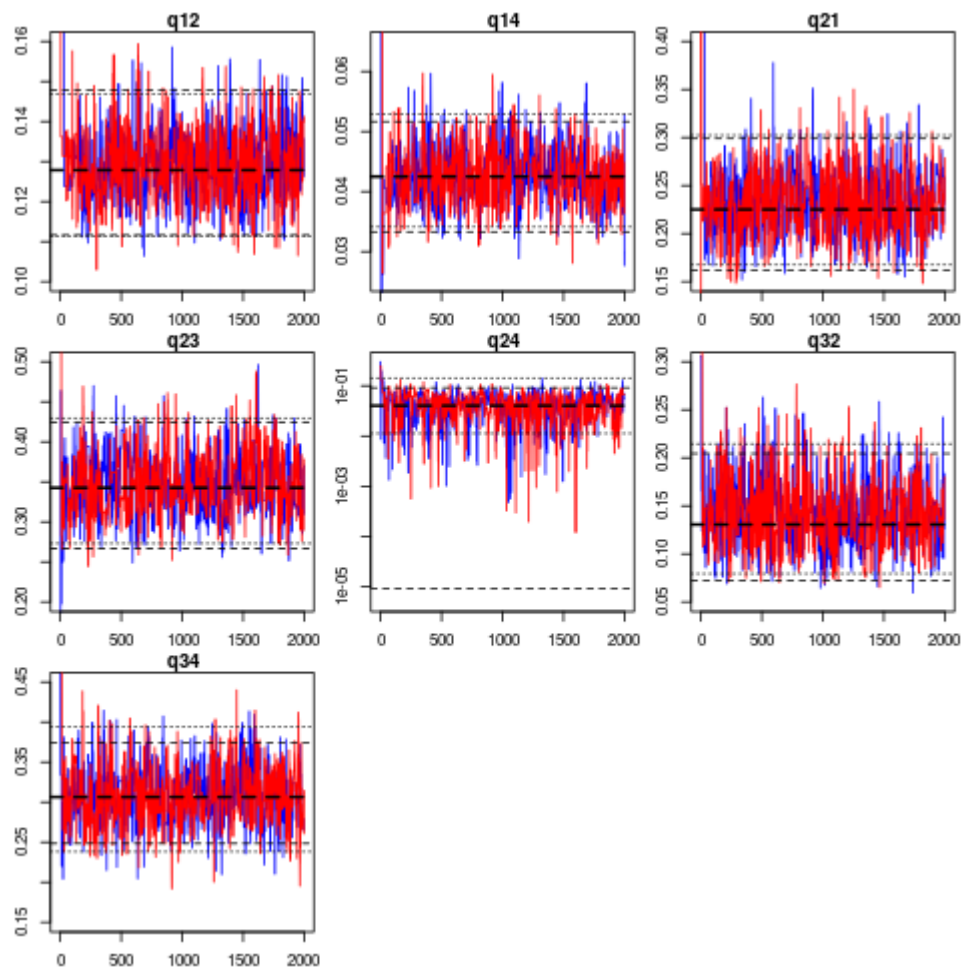


Figure S9. Parameter trajectories of two MCMC runs of the coronary allograft vasculopathy model with GNU MCSim (red and blue) and R *msm* estimates (MLE, thick dashed line; 95% confidence intervals, normal approximation, long dashes; 95% confidence intervals, bootstrapped, short dashes).

Section 6 – Model code and executable MCMC file

1. Model code

```
# Units
# Concentration:    Mol
# Time:             day
# Rates:            1/day

# States are the number of zebrafish embryo in each state, they can be output
States = {P_N_N,      # Probability to stay in no effect state
          P_E_E,      # ..... effects state
          P_H_H,      # ..... hatched state
          P_EH_EH,    # ..... hatched with effects state
          P_N_H,      # ..... go from normal to hatched state
          P_N_E,      # ..... effects state
          P_N_C,      # ..... coagulated state
          P_N_EH,     # ..... hatched with effects state
          P_E_C,      # ..... effects to coagulated state
          P_E_N,      # ..... effects to normal state
          P_E_H,      # ..... hatched state
          P_E_EH,     # ..... hatched with effects state
          P_H_EH,     # ..... hatched to hatched with effects state
          P_EH_H};    # ..... hatched with effects to hatched state

Inputs = {q_N_H, q_E_EH};
Outputs = {N_N_t,
           N_H_t,
           N_E_t,
           N_C_t,
           N_EH_t,
           Mass_Balance, # Mass balance of the model
           LL_N_N,
           LL_H_H,
           LL_EH_EH,
           LL_E_E,
           LL_N_E,
           LL_N_H,
           LL_N_C,
           LL_N_EH,
           LL_E_C,
           LL_E_N,
           LL_E_H,
           LL_E_EH,
           LL_H_EH,
           LL_EH_H,
           q_E_N,      # Transition rate from effects to normal state (/day)
           q_EH_H,     # Transition rate from hatched with effects to hatched (/day)
           q_H_EH,     # Transition rate from hatched to hatched with effects (/day)
           q_N_E,      # Transition rate from Normal to effects state (/day)
           q_N_C,      # Transition rate from Normal to coagulated state (/day)
           q_E_C};     # Transition rate from effects to coagulated state (/day)

# Parameters

# Concentration of chemical (M)
Dose;          # in fact, nominal medium concentration
Conc_t;        # time-varying concentration, will be calculated in ADDME.c
Conc_num = 21; # the line to read in concentration ratio input file

# Piecewise model parameters to fit the hatching rate
# Transition rate from Normal to hatched state (/day)
q_N_H_est;
# Transition rate from Effects to hatched with effects state (/day)
q_E_EH_est;
# Hatching time (day)
th;
```

```

# Statistical parameters
SD_qNH;
SD_qEEH;
SD_th;

# Transition rate from Normal to normal state (/day)
q_N_N;

# Transition rate from effects to effects state (/day)
q_E_E;

# Transition rate from hatched to hatched state (/day)
q_H_H;

# Transition rate from hatched with effects to hatched with effects state (/day)
q_EH_EH;

## Parameters needed to model relationships between transition rates and conc.
# q_E_N
qEN_zero;
k_qEN;

# q_EH_H
#qEHH_zero;
#k_qEHH;
u;
v;

# q_H_EH
qHEH_max;
qHEH_Hill;
qHEH_EC50;
#qHEH_zero;
#k_qHEH;

# q_N_E
qNE_zero;
k_qNE;

# q_N_C
qNC_zero;
k_qNC;

# q_E_C
qEC_zero;
k_qEC;

# Number of zebrafish embryo with no effect at time 0
N_N_t0;

# Number of zebrafish embryo with effects at time 0
N_H_t0;

# Number of zebrafish embryo with effects at time 0
N_E_t0;

# Number of zebrafish embryo coagulated at time 0
N_C_t0;

# Number of zebrafish embryo hatched with effects at time 0
N_EH_t0;

# Data: number of transitions
N_transitions;

```

```

#-----
# Dynamics
#-----

Dynamics {

    # Conc_t is time dependent and given by the PK model
    # We calculate it in the ADDME.c extension file:
    Inline( \#include "ADDME_v2.c" );

    # Transition rate from normal to effects state (/day)
    q_E_N = qEN_zero * exp(k_qEN * Conc_t);

    # Transition rate from hatched with effects to hatched state (/day)
    qEHH_zero = -u;          # donc entre 0 et 4
    k_qEHH      = v / 0.02; # donc entre -qEH_H_0 / 0.02 et 0, donc qEH,H > 0
    q_EH_H      = qEHH_zero + k_qEHH * Conc_t;

    # Transition rate from hatched to hatched with effects state (/day)
    q_H_EH = (qHEH_max * pow(Conc_t, qHEH_Hill)) /
              (pow(qHEH_EC50, qHEH_Hill) + pow(Conc_t, qHEH_Hill));
    #q_H_EH = qHEH_zero * exp(k_qHEH * Conc_t);

    # Transition rate from Normal to effects state (/day)
    q_N_E = qNE_zero * exp(k_qNE * Conc_t);

    # Transition rate from Normal to coagulated state (/day)
    q_N_C = qNC_zero + k_qNC * Conc_t;

    # Transition rate from effects to coagulated state (/day)
    q_E_C = qEC_zero + k_qEC * Conc_t;

    # Probability of zebrafish embryo to stay in no effect state
    dt(P_N_N) = - P_N_N * (q_N_H + q_N_E + q_N_C) + P_N_E * q_E_N;

    # Probability of zebrafish embryo to go from normal to hatched state
    dt(P_N_H) = P_N_N * q_N_H - P_N_H * q_H_EH + P_N_EH * q_EH_H;

    # Probability of zebrafish embryo to go from normal to effects state
    dt(P_N_E) = P_N_N * q_N_E - P_N_E * (q_E_N + q_E_C + q_E_EH);

    # Probability of zebrafish embryo to go from normal to coagulated state
    dt(P_N_C) = P_N_N * q_N_C + P_N_E * q_E_C;

    # Probability to go from normal to hatched with effects state
    dt(P_N_EH) = P_N_H * q_H_EH + P_N_E * q_E_EH - P_N_EH * q_EH_H;

    # Probability of zebrafish embryo to stay in hatched state
    dt(P_H_H) = - P_H_H * q_H_EH + P_H_EH * q_EH_H;

    # Probability to go from hatched to hatched with effects state
    dt(P_H_EH) = P_H_H * q_H_EH - P_H_EH * q_EH_H;

    # Probability of zebrafish embryo to go from effects to hatched state
    dt(P_E_H) = P_E_N * q_N_H - P_E_H * q_H_EH + P_E_EH * q_EH_H;

    # Probability of zebrafish embryo to go from effects to normal state
    dt(P_E_N) = - P_E_N * (q_N_H + q_N_E + q_N_C) + P_E_E * q_E_N;

    # Probability of zebrafish embryo to stay in effects state
    dt(P_E_E) = P_E_N * q_N_E - P_E_E * (q_E_N + q_E_C + q_E_EH);

    # Probability of zebrafish embryo to go from effects to coagulated state
    dt(P_E_C) = P_E_N * q_N_C + P_E_E * q_E_C;

    # Probability to go from effects to hatched with effects state
    dt(P_E_EH) = P_E_H * q_H_EH + P_E_E * q_E_EH - P_E_EH * q_EH_H;

```

```

# Probability to go from hatched with effects to hatched state
dt(P_EH_H) = - P_EH_H * q_H_EH + P_EH_EH * q_EH_H;

# Probability of zebrafish embryo to stay in hatched with effects state
dt(P_EH_EH) = P_EH_H * q_H_EH - P_EH_EH * q_EH_H;
}

#-----
# CalcOutputs
#-----

CalcOutputs {

# Probability to have N_x_t ZFE in state x at time t,
# if there are N_n_t0 ZFE in state n at time 0
N_N_t = N_N_t0 * P_N_N + N_E_t0 * P_E_N;
N_H_t = N_N_t0 * P_N_H + N_H_t0 * P_H_H + N_E_t0 * P_E_H + N_EH_t0 * P_EH_H;
N_E_t = N_N_t0 * P_N_E + N_E_t0 * P_E_E;
N_C_t = N_N_t0 * P_N_C + N_E_t0 * P_E_C + N_C_t0; # P_C_C = 1
N_EH_t = N_N_t0 * P_N_EH + N_H_t0 * P_H_EH + N_E_t0 * P_E_EH + N_EH_t0 * P_EH_EH;

# Mass balance
Mass_Balance = P_N_N + P_N_H + P_N_E + P_N_C + P_N_EH +
                P_E_C + P_E_N + P_E_EH + P_E_E + P_E_H +
                P_H_EH + P_H_H + P_EH_H + P_EH_EH;

# Likelihoods
LL_N_N = N_transitions * log((P_N_N > 0 ? P_N_N : 1E-30));
LL_N_H = N_transitions * log((P_N_H > 0 ? P_N_H : 1E-30));
LL_N_E = N_transitions * log((P_N_E > 0 ? P_N_E : 1E-30));
LL_N_C = N_transitions * log((P_N_C > 0 ? P_N_C : 1E-30));
LL_N_EH = N_transitions * log((P_N_EH > 0 ? P_N_EH : 1E-30));
LL_E_C = N_transitions * log((P_E_C > 0 ? P_E_C : 1E-30));
LL_E_N = N_transitions * log((P_E_N > 0 ? P_E_N : 1E-30));
LL_E_EH = N_transitions * log((P_E_EH > 0 ? P_E_EH : 1E-30));
LL_E_E = N_transitions * log((P_E_E > 0 ? P_E_E : 1E-30));
LL_E_H = N_transitions * log((P_E_H > 0 ? P_E_H : 1E-30));
LL_H_EH = N_transitions * log((P_H_EH > 0 ? P_H_EH : 1E-30));
LL_H_H = N_transitions * log((P_H_H > 0 ? P_H_H : 1E-30));
LL_EH_H = N_transitions * log((P_EH_H > 0 ? P_EH_H : 1E-30));
LL_EH_EH = N_transitions * log((P_EH_EH > 0 ? P_EH_EH : 1E-30));

} # End of output calculation
End.

```

2. PK link file

```

/* file containing code to read a concentration vs time profile and
   interpolate it.*/

#define NCONCS 8
#define NTIMES 121

static int firstcall = 1;
static double C_read[NCONCS][NTIMES];
FILE* fp;
int i, j;
int Cindex, Tindex;
double Conc_t1, Conc_t2;

if (firstcall == 1) { /* on first call, read the concentration file */
    firstcall = 0; /* don't do it again */
    if (!(fp = fopen("VPA_internal.txt", "r"))) {
        printf("Error: file VPA_internal.txt not found - Exiting.\n\n");
        exit(0);
    }
    for (j = 0; j < NTIMES; j++) {

```

```

    C_read[0][j] = 0; /* at expoure zero, internal dose is always 0 */
}
for (i = 1; i < NCONCS; i++) {
    printf("reading concentration profiles #%d\n", i+1);
    for (j = 0; j < NTIMES; j++) {
        if ( fscanf(fp, "%lg", &(C_read[i][j])) == EOF ) {
            printf ("Error: VPA_internal.txt file too short - Exiting\n\n");
            exit(0);
        }
        else { /* reading ok */
            /* printf("C[%d,%d] = %g\n", i, j, C_read[i][j]); */
        }
    }
} /* for i */
} /* end of on first call */

/* Find the position of current Dose in the official list of 8 doses */
double Ref_concs[NCONCS] = {0.000, 6.25E-6, 12.5E-6, 25.E-6,
                             50E-6, 100.E-6, 200.E-6, 400E-6};

for (i = 1; i < NCONCS; i++) {
    if (Dose < Ref_concs[i]) break;
}
Cindex = i-1;

/* Find the position of current time in the C_read array.
   There are 121 times from 0 to 5 days by increments of 1/24.0 */
Tindex = (int) ((*pdTime) * 24);

/* Compute concentration at current time by interpolation on the lower bound
   of Concentration */
if (Tindex < (NTIMES - 1))
    Conc_t1 = C_read[Cindex][Tindex] +
              (C_read[Cindex][Tindex + 1] - C_read[Cindex][Tindex]) *
              ((*pdTime) * 24 - Tindex);
else /* at time boundary */
    Conc_t1 = C_read[Cindex][Tindex];

/* Compute concentration at current time by interpolation on the upper bound
   of Concentration */
if (Cindex < (NCONCS - 1)) {
    if (Tindex < (NTIMES - 1))
        Conc_t2 = C_read[Cindex+1][Tindex] +
                  (C_read[Cindex+1][Tindex + 1] - C_read[Cindex+1][Tindex]) *
                  ((*pdTime) * 24 - Tindex);
    else /* at time boundary */
        Conc_t2 = C_read[Cindex+1][Tindex];
}
else { /* at concentration boundary */
    if (Tindex < (NTIMES - 1))
        Conc_t2 = C_read[Cindex][Tindex] +
                  (C_read[Cindex][Tindex + 1] - C_read[Cindex][Tindex]) *
                  ((*pdTime) * 24 - Tindex);
    else /* time boundary */
        Conc_t2 = C_read[Cindex][Tindex];
}

/* Interpolate between the two concentrations calculated above; Conc_t will be
   used in the multistate model */
if (Cindex < (NCONCS - 1)) {
    Conc_t = Conc_t1 + (Conc_t2 - Conc_t1) *
                (Dose - Ref_concs[Cindex]) /
                (Ref_concs[Cindex+1] - Ref_concs[Cindex]);
}
else {
    Conc_t = Conc_t1;
}
} /* End */

```

Titre : Modélisation pharmacocinétique et pharmacodynamique de la toxicité développementale de l'embryon de poisson zèbre

Mots-clés : Danio rerio, Pharmacocinétique, Pharmacodynamique, Modélisation, Développement, Malformation

Résumé : En évaluation de risques, comprendre et prédire les effets des substances chimiques sur le développement des êtres vivants sont d'une nécessité grandissante pour évaluer la toxicité environnementale ou thérapeutique. Au vu de ses nombreux avantages expérimentaux, réglementaires et de ses homologies génétiques avec l'Homme, l'embryon de poisson zèbre est largement utilisé pour les tests de toxicité. Dans le cadre du projet européen *Eu-ToxRisk* et pour améliorer l'extrapolation à l'Homme des données obtenues chez l'embryon de poisson zèbre, l'objectif de nos travaux était de développer deux modèles mathématiques pour expliquer et prédire les malformations observées au cours des cinq premiers jours de développement de l'embryon de poisson zèbre, suite à l'exposition à différentes concentrations d'acide valproïque (VPA), tératogène notoire, et neuf de ses analogues : Premièrement, un modèle pharmacocinétique (PK) générique basé sur la physiologie (PBPK) capable de prédire les concentrations de composés neutres ou ionisables dans dix organes d'intérêt d'étude de l'embryon. La croissance des organes est considérée et les coefficients de partage entre milieu de culture et organes sont prédits par un sous-modèle QSAR. La clairance métabolique et un facteur de correction des coefficients de partage ont été estimés simultanément par ajustement bayésien du modèle aux données de concentrations de VPA et de ses analogues dans l'embryon et le milieu de culture au cours de cinq jours de dosages répétés. Deuxièmement, un modèle multi-états Markovien pharmacodynamique (PD) capable de prédire l'occurrence de malformations en fonction du temps et de la concentration d'exposition en VPA. Enfin, pour décrire le comportement du VPA dans l'organisme de l'embryon et l'impact de l'exposition au cours du temps sur le développement de malformations, les deux modèles ont été couplés pour générer un modèle PK-PD. La concentration interne moyenne de VPA, prédite par le modèle PBPK, contrôle les transitions aléatoires entre états. Par ajustement bayésien aux données de malformations, les taux de transitions d'un état à l'autre ont été estimés, permettant ainsi de déterminer les probabilités de transition résultantes et les concentrations effectives associées.

Title: Pharmacokinetic and pharmacodynamic modeling of the developmental toxicity of the zebrafish embryo

Keywords: Danio rerio, Pharmacokinetics, Pharmacodynamics, Modelling, Development, Malformation

Abstract: In risk assessment, understanding and predicting the effects of chemical compounds on the development of living beings are needed to understand and manage environmental and therapeutic toxicity. Due to its many experimental and regulatory advantages, and genetic homologies with humans, the zebrafish embryo is widely used for toxicity testing. As part of the European project *Eu-ToxRisk*, the objective of this work was to improve the extrapolation of the zebrafish embryo data to humans by developing two mathematical models able to explain and predict the occurrence of malformations during the zebrafish embryo first five days of development, following exposure valproic acid (VPA), a notorious teratogen, and nine of its analogues. First, a generic physiology-based pharmacokinetic model (PBPK) able to predict the concentrations of neutral or ionizable compounds in ten embryonic organs of interest. Organ growth is considered and the culture medium to organ partition coefficients are predicted by a QSAR sub-model. Metabolic clearance and a partition coefficient correction factor were simultaneously estimated by Bayesian calibration of the model with embryo and culture medium concentration data of VPA and its analogues over five days of repeated dosing. Second, a multi-state Markovian pharmacodynamic (PD) model to predict the occurrence of malformations as a function of time and VPA exposure concentration. Finally, to describe the behavior of VPA in the embryo and the impact of time and exposure on malformations, the two models were coupled to generate a PK-PD model. The average internal concentration of VPA, predicted by the PBPK model, controls random transitions between states. By Bayesian calibration with malformation data, the transition rates between states were estimated, allowing estimation of transition probabilities and associated effective concentrations.