

UNIVERSITE TOULOUSE III PAUL SABATIER
FACULTE DE SANTE
DEPARTEMENT DES SCIENCES PHARMACEUTIQUES

ANNEE : 2024

THESE 2024 / TOU3 / 2045

THESE

POUR LE DIPLOME D'ETAT DE DOCTEUR EN PHARMACIE

Présentée et soutenue publiquement

par

DE KETELAERE VALENTINE

**LA PERTINENCE DES DONNÉES FOURNIES PAR LES
MÉTHODES ALTERNATIVES À L'EXPERIMENTATION ANIMALE
DANS L'ÉVALUATION DU DANGER EN TOXICOLOGIE : ÉTAT
DES LIEUX, ÉTUDES DE CAS ET PERSPECTIVES.**

Le 05 Juillet 2024

Directrice de thèse : Pr Boutet Élisabeth

JURY

Présidente : Pr Deraeve Céline
1er assesseur : Pr Boutet Élisabeth
2ème assesseur : Dr Sotton Benoît

LA PERTINENCE DES DONNÉES FOURNIES PAR LES MÉTHODES ALTERNATIVES À L'EXPERIMENTATION ANIMALE DANS L'ÉVALUATION DU DANGER EN TOXICOLOGIE : ÉTAT DES LIEUX, ÉTUDES DE CAS ET PERSPECTIVES.

RESUME en français

Les méthodes traditionnelles d'évaluation du danger en toxicologie, effectuées sur les animaux, se voient depuis plusieurs décennies remises en question pour des raisons éthiques, pratiques et économiques, favorisant l'émergence de méthodes alternatives. Ces dernières, également appelées « Nouvelles Approches Méthodologiques » (NAM) en toxicologie, incluent majoritairement des modélisations mathématiques (*in silico*) et des tests sur les cellules (*in vitro*). Le présent document a pour objectif d'en évaluer la justesse, en prenant pour exemple une étude de cas sur un critère d'évaluation particulier, la toxicité orale aiguë. La méthode retenue consiste à comparer les valeurs expérimentales obtenues sur les animaux pour différents groupes de substances avec les données obtenues à l'aide de deux approches NAM : les read-across et la méthode QSAR (« Quantitative Structure-Activity Relationship »). Cette comparaison vise à déterminer si les approches NAM influencent l'évaluation des risques, et donc la classification selon les critères « CLP » (« Classification, Labelling and Packaging ») propre à la toxicité aiguë, en vue de l'élaboration de directives d'essai par les autorités européennes. Cette étude permet de montrer la pertinence croissante des NAM, qui devrait s'intensifier avec les progrès effectués dans le domaine de l'intelligence artificielle (IA). Cette dernière offre la possibilité d'appréhender la complexité des ensembles de données liées à la toxicité, et peut être appliquée à de nombreux aspects de la toxicologie prédictive, contribuant ainsi à la réduction de l'expérimentation animale. Malgré des avantages prometteurs, des défis persistent quant à l'adoption de ces nouvelles méthodes par les autorités réglementaires. Elles devront faire l'objet de futures collaborations scientifiques et répondre à des exigences éthiques et réglementaires. Ces nouvelles méthodes alternatives, appliquées à l'évaluation du danger en toxicologie, pourraient significativement transformer la réglementation entourant les substances chimiques.

Titre et résumé en Anglais : voir en quatrième de couverture

DISCIPLINE administrative : Toxicologie

MOTS-CLES : Toxicologie, évaluation des dangers, risques, substances chimiques, *in silico*, nouvelles approches méthodologiques, expérimentation animale, éthique, intelligence artificielle, modèles prédictifs.

Directrice de thèse : Pr Boutet Elisa

PERSONNEL ENSEIGNANT
du Département des Sciences Pharmaceutiques
de la Faculté de santé
au 08 mars 2023

Professeurs Emérites

Mme BARRE A.	Biologie Cellulaire
M. BENOIST H.	Immunologie
Mme ROQUES C.	Bactériologie - Virologie
M. ROUGE P.	Biologie Cellulaire
M. SALLES B.	Toxicologie

Professeurs des Universités

Hospitalo-Universitaires

Mme AYYOUB M.	Immunologie
M. CESTAC P.	Pharmacie Clinique
M. CHATELUT E.	Pharmacologie
Mme DE MAS MANSAT V.	Hématologie
M. FAVRE G.	Biochimie
Mme GANDIA P.	Pharmacologie
M. PARINI A.	Physiologie
M. PASQUIER C.	Bactériologie - Virologie
Mme ROUSSIN A.	Pharmacologie
Mme SALLERIN B. (Directrice-adjointe)	Pharmacie Clinique
M. VALENTIN A.	Parasitologie

Universitaires

Mme BERNARDES-GENISSON V.	Chimie thérapeutique
Mme BOUTET E.	Toxicologie - Sémiologie
Mme COSTE A.	Parasitologie
Mme COUDERC B.	Biochimie
M. CUSSAC D. (Doyen-directeur)	Physiologie
Mme DERA EVE C.	Chimie Thérapeutique
M. FABRE N.	Pharmacognosie
Mme GIROD-FULLANA S.	Pharmacie Galénique
M. GUIARD B.	Pharmacologie
M. LETISSE F.	Chimie pharmaceutique
Mme MULLER-STAU MONT C.	Toxicologie - Sémiologie
Mme REYBIER-VUATTOUX K.	Chimie analytique
M. SEGUI B.	Biologie Cellulaire
Mme SIXOU S.	Biochimie
Mme TABOULET F.	Droit Pharmaceutique
Mme WHITE-KONING M.	Mathématiques

Maîtres de Conférences des Universités

Hospitalo-Universitaires		Universitaires	
M. DELCOURT N.	Biochimie	Mme ARELLANO C. (*)	Chimie Thérapeutique
Mme JUILLARD-CONDAT B.	Droit Pharmaceutique	Mme AUTHIER H.	Parasitologie
Mme KELLER L.	Biochimie	M. BERGE M. (*)	Bactériologie - Virologie
M. PUISSET F.	Pharmacie Clinique	Mme BON C. (*)	Biophysique
Mme ROUCH L.	Pharmacie Clinique	M. BOUAJILA J. (*)	Chimie Analytique
Mme ROUZAUD-LABORDE C	Pharmacie Clinique	M. BROUILLET F.	Pharmacie Galénique
Mme SALABERT A.S.	Biophysique	Mme CABOU C.	Physiologie
Mme SERONIE-VIVIEN S (*)	Biochimie	Mme CAZALBOU S. (*)	Pharmacie Galénique
Mme THOMAS F. (*)	Pharmacologie	Mme CHAPUY-REGAUD S. (*)	Bactériologie - Virologie
		Mme COLACIOS C. (*)	Immunologie
		Mme ECHINARD-DOUIN V. (*)	Physiologie
		Mme EL GARAH F.	Chimie Pharmaceutique
		Mme EL HAGE S.	Chimie Pharmaceutique
		Mme FALLONE F.	Toxicologie
		Mme FERNANDEZ-VIDAL A.	Toxicologie
		Mme GADEA A.	Pharmacognosie
		Mme HALOVA-LAJOIE B.	Chimie Pharmaceutique
		Mme JOUANJUS E.	Pharmacologie
		Mme LAJOIE-MAZENC I.	Biochimie
		Mme LEFEVRE L.	Physiologie
		Mme LE LAMER A-C. (*)	Pharmacognosie
		M. LE NAOUR A.	Toxicologie
		M. LEMARIE A.	Biochimie
		M. MARTI G.	Pharmacognosie
		Mme MONFERRAN S	Biochimie
		M. PILLOUX L.	Microbiologie
		M. SAINTE-MARIE Y.	Physiologie
		M. STIGLIANI J-L.	Chimie Pharmaceutique
		M. SUDOR J. (*)	Chimie Analytique
		Mme TERRISSE A-D.	Hématologie
		Mme TOURRETTE-DIALLO A. (*)	Pharmacie Galénique
		Mme VANSTEEELANDT M.	Pharmacognosie

(*) Titulaire de l'habilitation à diriger des recherches (HDR)

Enseignants non titulaires

Assistants Hospitalo-Universitaires		Attaché Temporaire d'Enseignement et de Recherche (ATER)	
M. AL SAATI A	Biochimie	Mme HAMZA Eya	Biochimie
Mme BAKLOUTI S.	Pharmacologie	Mme MALLI Sophia	Pharmacie Galénique
Mme CLARAZ P.	Pharmacie Clinique	M. TABTI Redouane	Chimie Thérapeutique
Mme CHAGNEAU C.	Microbiologie		
Mme DINTILHAC A.	Droit Pharmaceutique		
M. LE LOUEDEC F.	Pharmacologie		
Mme RIGOLOT L.	Biologie Cellulaire, Immunologie		
Mme STRUMIA M.	Pharmacie Clinique		

Remerciements

A l'ensemble des équipes pédagogiques,

En premier lieu à la Faculté des Sciences Pharmaceutiques de Toulouse, pour leur accompagnement quotidien au cours de l'ensemble de mes années d'études, et pour leur volonté de former des professionnels de santé compétents. Je tiens également à remercier les enseignants de l'Université Paris Cité, tout particulièrement ceux du Master 2 THERV que j'ai suivi pendant un an. L'élaboration de cette thèse a été fondée sur la qualité de leurs enseignements dispensés en toxicologie.

A ma directrice de thèse, Madame la Professeure Elisa Boutet,

Pour son engagement, son temps, son écoute et sa bienveillance depuis trois ans. Nous nous sommes rencontrées au sein du « sous-parcours recherche », réservé aux étudiants intéressés par la découverte scientifique. Vous n'avez eu de cesse de faire fructifier ce parcours de la filière Industrie, favorisant l'accès au monde de la recherche par un accompagnement personnalisé. Ce dernier a pris toute son ampleur quand, face aux difficultés que j'ai rencontrées, vous avez su m'épauler de la plus altruiste des façons. Merci d'avoir accepté de devenir la directrice de ma thèse d'exercice. Je n'aurais pu trouver de meilleure encadrante, dont les compétences en toxicologie égalent les valeurs humaines. Veuillez trouver dans ces lignes un bref aperçu de ma profonde gratitude.

Au Docteur Benoit Sotton,

Pour avoir accepté de prendre part à ce jury. Une deuxième année à superviser ma production, une nouvelle fois remercié, mais jamais suffisamment à mon goût. Merci ta bienveillance, ta disponibilité et ton enthousiasme pour ce sujet (quitte à jouer avec QSAR Toolbox un peu trop tard le soir). Faisons un pas de côté : merci aussi pour les conseils professionnels autant que personnels, pour les quelques (nombreux) verres partagés, et les très pointues recommandations musicales. Je ne pourrais être trop formelle en m'adressant à toi, car au-delà d'un maître de stage hors-pair, j'ai fait la connaissance d'un ami.

A ma présidente de jury, la Professeure Céline Deraeve,

Pour m'avoir fait l'honneur de présider cette thèse. Veuillez trouver ici l'expression de ma profonde reconnaissance.

A mes parents,

Enfin, terminé ! Maman, Papa, vous allez pouvoir souffler. Plus de révisions panique en fin d'année, fini les « j'peux pas j'dois bosser ». Sept ans de pharmacie, c'est long ; ça n'a pas dû être facile tous les jours, et je ne saurais comment vous remercier. « Normal pour des parents », direz-vous, mais merci de m'avoir permis de réaliser les études que je souhaitais, et pour votre accompagnement dans leur réussite. Merci d'être venus si souvent me chercher à la BU au moment du concours (avec un dîner tout prêt), puis, dans les années qui ont suivi, d'avoir pris un temps conséquent à relire chacun de travaux. Vous m'avez écouté apprendre, répéter, m'énerver, désespérer parfois, mais vous avez toujours été là, vous adaptant à une partie de la science que vous maîtrisiez peu. Maman, modèle de carrière professionnelle autant que de vie personnelle, merci de m'avoir inculqué la rigueur autant que la tendresse. Papa, merci pour ta patience, ton aide infaillible, pour ta sensibilité à peine dissimulée.

L'estime que vous m'accordez, ainsi que votre soutien sans faille dans tous les projets que j'entreprends sont mon moteur de tous les jours.

A petite sœur Juliette,

Pour vingt-trois années communes, de bas parfois mais surtout de hauts, de complicité sans limite. Merci pour ta patience, tes conseils avisés (malgré des appels qui peuvent s'éterniser), ta droiture et ton sens de la famille. Tu es et resteras la personne la plus créative que je connaisse et tu détiens plus de ressources que tu ne veux l'admettre. Comme promis, on a commencé par partager un canapé-lit à Paris mais on finira par mutuellement « se mettre à l'abri ».

A mon petit frère Raphaël,

A ma demi-portion qui a beaucoup trop grandi. Merci pour les conversations toujours intéressantes que nous partageons malgré la différence d'âge, tes avis stylistiques bien tranchés, ta positivité. Je suis fière du jeune homme plein d'ambition que tu es devenu.

A tous les membres de ma famille,

Si nombreux pour mon plus grand bonheur : je ne pourrais en citer tous les noms. Merci pour votre affection et votre soutien inconditionnel.

A mes ami.e.s,

Les premier.ère.s, qui datent d'une période encore peu affectée par les aléas étudiants. Jeanne, pour toutes ces années à tes côtés, de la classe d'anglais de 6^{ème} (je reste persuadée que tu étais la nouvelle correspondante) aux cafés côte-à-côte en terrasse parisienne, nous avons vécu trop de rebondissements pour qu'ils soient énumérés ici. Julie, l'ambassadrice, toujours aussi présente malgré les milliers de kilomètres qui nous séparent. (Tata) Rapha, modèle de courage et de persévérance, merci pour cette belle amitié. Aminata, merci d'avoir donné une seconde chance à la capitale, j'ai hâte des souvenirs qui nous attendent. Les années lycées ne vont pas non plus sans Josi, dont la joie de vivre m'a suivi même dans les révisions de première année, mais aussi Victor et Romane, piliers du « cercle ».

Sont arrivé.e.s ensuite celles et ceux de la fac, certain.e.s dès le début, d'autres plus tard, symboles de la longueur mais aussi de la richesse de ces études. Merci à l'infatigable groupe des GINTO, depuis notre rencontre au week-end d'inté à la consolidation du groupe, à travers de nombreux voyages et malgré la distance qui nous sépare aujourd'hui, la coloc ne s'arrêtera jamais. Amélie, merci d'être une amie si fidèle, de regorger de nouvelles idées (et de nouveaux surnoms), et de m'avoir fait rire aux larmes plus d'une fois. Clara, merci de toujours répondre présente, de trouver les mots, merci pour tes conseils précieux. Mathilde, en attendant les vagues qui restent à surfer, merci pour ta générosité et ton optimisme au quotidien. En pharma, j'ai aussi croisé la route de Clacul (Minaj), qui a emmené de franches rigolades et du folklore aveyronnais, et de Cam G (Ayebo), togolaise dans l'âme, merci pour ces expériences associatives. J'ai ensuite rencontré Maëlle, « dancing queen », Olivier, notre vedette locale, Manon, pleine de douceur, et Clara, conseillère artistique de talent. Des noms qui me rappellent les souvenirs inoubliables que vous m'avaient offert. Enfin, aux Tox Babes : Eugénie, Inès, Louise, Laure, Clémentine, merci de m'avoir incluse dans votre groupe au sein du master ; mon expérience dans la capitale n'aurait pas été aussi belle sans vous.

Au programme Erasmus,

Pour m'avoir offert la chance de rencontrer de belles personnes à Newcastle. Masé, Hanna, Ysanne, Valentin, Fine, Zoe, Juan, Peter, Nicola, ... merci à tous pour ces souvenirs, ces amitiés solides construites en dépit de la barrière de la langue. Merci également à Daniel, Mélanie et Sarah, pour avoir été les meilleurs colocs bruxellois.

Aux « EcoMundiens »,

Qui ont ensoleillé les six mois de stage à l'origine de ce sujet de thèse. Merci à Juliette, pour m'avoir fait autant confiance (promis on ira voir du punk), à Djamil, pour les sets et les commentaires sympas, à Sofia pour les précieux conseils. Je ne peux oublier le « Nuevo Mundo », avec mes fidèles acolytes stagias Caro et Clara, et les non moins stagias Diego et Audrey : merci pour les bons moments partagés. Merci à tous ceux de l'équipe REACH/biocide, ainsi qu'aux autres équipes, qui ont contribué de près ou de loin à l'élaboration de cette thèse, et qui ont dû subir de nombreuses présentations d'entraînement.

A Augustin,

Gus à l'occasion, Gugus souvent. De flirt « expérimental » à amoureux émérite, à travers la Belgique, l'Italie et la France, tu as réussi à me supporter, à toujours être présent malgré la distance. Toujours attentionné et sincère ; merci pour ton aide précieuse, de la conception aux dernières corrections de la thèse, pour les lignes de code et la livraison de chocolat, merci pour tout l'amour que tu me portes.

Enfin, à ceux qui ne sont plus,

Emportés au moment où je planchais sur ce sujet de thèse. A toi, Papi Michel, grand-père épatant, tu étais à l'écoute autant que ton handicap te le permettait, toujours digne. Papi Jacques, co-fondateur d'une belle et grande famille, d'une gentillesse et d'une bonne humeur à toute épreuve. Merci pour vos valeurs, pour avoir contribué à faire de moi la femme que je suis aujourd'hui. Je ne vous oublie pas.

Grâce à chacune et chacun de vous, rédiger cette thèse n'aura pas été un travail solitaire.

« Rien n'est plus puissant qu'une idée dont l'heure est venue »

- Victor Hugo

Table des matières

Remerciements	8
Résumé	16
Liste des illustrations	18
A. Figures	18
B. Tableaux	19
Liste des abréviations	20
Préambule	22
I. Evolution des méthodes traditionnelles d'expérimentation en toxicologie vers les méthodes alternatives	25
A. État de l'art de l'expérimentation animale en évaluation des risques	25
B. Les défaillances de l'expérimentation animale traditionnelle	28
1. Principales difficultés	28
2. Limites éthiques	31
3. Développement nécessaire des méthodes alternatives	33
C. Émergence des « Nouvelles Approches Méthodologiques » (NAM) en toxicologie	35
1. Définitions et principes	35
2. Types de NAM	39
3. Avantages et inconvénients des différentes approches	44
D. Cadre réglementaire et intégration des NAM dans l'évaluation des risques	45
1. Réglementations actuelles	45
2. Initiatives et directives réglementaires en faveur des NAM	48
3. Défis et perspectives pour une meilleure intégration	50
II. Évaluation de la pertinence des données NAM dans l'approche d'évaluation du danger : études de cas sur la toxicologie orale aiguë	53
A. Introduction	53
1. Contexte	53
2. Définition des méthodes utilisées	53
3. Toxicité orale aiguë	56
4. Objectifs de recherche	58
B. Matériel et méthodes	60
1. Outils utilisés	60
2. Approches	64

C.	Résultats	69
1.	Étude de cas 1 : Carbamates	69
2.	Étude de cas n° 2 : Phosphorothioates	76
3.	Étude de cas n° 3 : Esters organophosphorés.....	81
4.	Total des prédictions exactes.....	86
D.	Discussion.....	87
E.	Résumé des caractéristiques des substances	93
1.	Étude de cas 1.....	93
2.	Étude de cas 2.....	94
3.	Étude de cas n° 3.....	96
III.	L'intelligence artificielle au service de la prédiction toxicologique	97
A.	Introduction.....	97
B.	Fondements de l'intelligence artificielle en toxicologie	99
C.	Applications de l'intelligence artificielle en pharmacologie et toxicologie prédictive 104	
1.	Prédiction de la cardiotoxicité par blocage du canal potassique (gène hERG). 104	
2.	Prédiction des atteintes hépatiques d'origine médicamenteuse (DILI).....	106
3.	Prédiction de l'inhibition de l'acétylcholinestérase (AChE).....	108
D.	Applications de l'intelligence artificielle en pharmacovigilance.....	110
E.	Intelligence artificielle : enjeux éthiques et réglementaires.....	115
F.	Défis, limites et perspectives futures	121
	Conclusion	125
	Bibliographie	130

Résumé

Les méthodes traditionnelles d'évaluation du danger en toxicologie, effectuées sur les animaux, se voient depuis plusieurs décennies remises en question pour des raisons éthiques, pratiques et économiques, favorisant l'émergence de méthodes alternatives. Ces dernières, également appelées « Nouvelles Approches Méthodologiques » (NAM) en toxicologie, incluent majoritairement des modélisations mathématiques (*in silico*) et des tests sur les cellules (*in vitro*). Le présent document a pour objectif d'en évaluer la justesse, en prenant pour exemple une étude de cas sur un critère d'évaluation particulier, la toxicité orale aiguë. La méthode retenue consiste à comparer les valeurs expérimentales obtenues sur les animaux pour différents groupes de substances avec les données obtenues à l'aide de deux approches NAM : les read-across et la méthode QSAR (« Quantitative Structure-Activity Relationship »). Cette comparaison vise à déterminer si les approches NAM influencent l'évaluation des risques, et donc la classification selon les critères « CLP » (« Classification, Labelling and Packaging ») propre à la toxicité aiguë, en vue de l'élaboration de directives d'essai par les autorités européennes. Cette étude permet de montrer la pertinence croissante des NAM, qui devrait s'intensifier avec les progrès effectués dans le domaine de l'intelligence artificielle (IA). Cette dernière offre la possibilité d'appréhender la complexité des ensembles de données liées à la toxicité, et peut être appliquée à de nombreux aspects de la toxicologie prédictive, contribuant ainsi à la réduction de l'expérimentation animale. Malgré des avantages prometteurs, des défis persistent quant à l'adoption de ces nouvelles méthodes par les autorités réglementaires. Elles devront faire l'objet de futures collaborations scientifiques et répondre à des exigences éthiques et réglementaires. Ces nouvelles méthodes alternatives, appliquées à l'évaluation du danger en toxicologie, pourraient significativement transformer la réglementation entourant les substances chimiques.

Liste des illustrations

A. Figures

Figure 1: Bénéficiaires des expérimentations animales en France	27
Figure 2: Schéma général des composantes des NAM.....	36
Figure 3: Schéma conceptuel des principales caractéristiques d'un AOP.....	38
Figure 4: Liste non exhaustive de méthodes, outils et applications NAM.	40
Figure 5: Schéma de la prédiction et de l'évaluation de la toxicité orale aiguë à l'aide d'outils <i>in silico</i>	59
Figure 6 : Étapes conduisant à la prédiction d'une donnée, sous QSAR Toolbox.	61
Figure 7 : Les différents composants de l'interface de QSAR Toolbox.....	61
Figure 8: Structure générale des carbamates	70
Figure 9: Capture d'écran de QSAR Toolbox montrant le début et la fin du jeu de données avec les informations d'identification chimique et la structure des membres du groupe des carbamates.....	70
Figure 10: Capture d'écran de QSAR Toolbox montrant les résultats du profilage des 4 premières substances du groupe.....	71
Figure 11: Capture d'écran d'une prédiction par read-across, pour une substance chimique du groupe..	74
Figure 12: Résumé des performances de prédiction de DL ₅₀ en pourcentages, pour cette première étude de cas comprenant 33 substances chimiques.....	75
Figure 13 Performance en termes de prédiction pour les trois approches par catégorie, comparée aux valeurs expérimentales.....	76
Figure 14: Capture d'écran de QSAR Toolbox montrant le début et la fin de la matrice de données avec les informations d'identification chimique et la structure des membres du groupe des phosphorothioates.....	77
Figure 15: Structure générale des phosphorothioates.....	77
Figure 16: Résumé de performances en termes de prédiction du DL ₅₀ en pourcentages, pour la deuxième étude de cas comprenant 29 substances chimiques.....	79
Figure 17: Performance de prédiction selon les trois approches, par catégorie.	81
Figure 18: Structure générale des esters organophosphorés	82

Figure 19: Capture d'écran de QSAR Toolbox montrant le début et la fin de la matrice de données, avec les informations d'identification chimique et la structure des membres de ce groupe d'esters organophosphorés.....	82
Figure 20: Performance des prédictions selon les trois approches.	84
Figure 21: Performance de prédiction selon les trois approches, par catégorie..	85
Figure 22 : Schéma de construction d'un modèle de « machine learning » (ML).....	101
Figure 23 : Enjeux éthiques soulevés au cours du cycle de vie des données	116

B. Tableaux

Table 1: Total des prédictions.	86
Table 2: Catégories CLP de toxicité.....	88
Table 3: Signalement des dangers par étiquetage pour la toxicité orale aiguë..	90

Liste des abréviations

AchE - Acétylcholinestérase	IRM - Imagerie à Résonance Magnétique
ADME - Absorption, Distribution, Métabolisation et Elimination	IUPAC - <i>International Union of Pure and Applied Chemistry</i>
ANN - <i>Artificial Neural Network</i>	IVIVE - <i>in vitro-in vivo extrapolation</i>
ANSM - Agence Nationale de Sécurité du Médicament et des Produits de Santé	JCR - <i>Journal Citation Report</i>
AOP - <i>Adverse Outcome Pathway</i>	MIE - <i>Molecular Initiating Event</i>
AOT DB - <i>Acute Oral Toxicity DataBase</i>	NAM - Nouvelles Approches Méthodologies
CAS - <i>Chemical Abstracts Service</i>	NIH - <i>National Institute of Health</i>
CLP - <i>Classification, Labelling and Packaging</i>	NLP - <i>Natural Language Processing</i>
CRPV - Centre Régional de Pharmacovigilance	OASIS - <i>Optimized Approach based on Structural Indices Set</i>
DILI - <i>Drug Induced Liver Injury</i>	OCDE - Organisation de Coopération et de Développement Economiques
DL ₅₀ - Dose Létale 50	OPE - <i>Organophosphate Ester</i>
DNN - <i>Deep Neural Network</i>	PBK - <i>Physiologically Based Kinetics</i>
DPRA - <i>Direct Peptide Reactivity Assay</i>	QSAR - <i>Quantitative Structure Activity Relationship</i>
DTU - <i>Technical University of Denmark</i>	RASAR - <i>Read-Across Structure Activity Relationship</i>
ECHA - <i>European Chemicals Agency</i>	REACH - <i>Registration, Evaluation, Authorisation and Restriction of Chemicals</i>
EFSA - <i>European Food Safety Authority</i>	RF - <i>Random Forest</i>
EPA - <i>Environmental Protection Agency</i>	RI - <i>Reliability Index</i>
FDA - <i>Food and Drug Administration</i>	TTC - <i>Threshold of Toxicological Concern</i>
GHS - <i>Globally Harmonized System</i>	UE - Union Européenne
hERG - <i>Human Ether-a-go-go</i>	
IA - Intelligence Artificielle	
ICE - <i>Integrated Chemical Environment</i>	

Préambule

Contexte de la toxicologie et de l'évaluation des risques

L'évaluation du risque toxicologique repose sur deux piliers fondamentaux : le danger inhérent à la substance chimique étudiée et l'exposition à cette substance. Le danger d'une substance est une caractéristique propre à cette substance, ce qui signifie qu'il dépend uniquement de ses propriétés intrinsèques et n'est pas influencé par des facteurs externes. Il concerne des aspects biologiques tels que la cancérogénicité, la néphrotoxicité, l'hépatotoxicité, la sensibilisation cutanée, etc. Ces caractéristiques sont liées à la structure de la molécule et à ses propriétés, elles existent indépendamment de la présence ou de la contamination d'un individu. L'exposition à la substance, quant à elle, se réfère à la quantité de substance chimique à laquelle un être humain est exposé, la durée et la voie d'exposition (alimentation, contact cutané, inhalation, etc.). Cette exposition dépend uniquement de l'environnement : consommation, lieux de résidence ou produits chimiques utilisés dans la vie quotidienne. Enfin, le risque représente la probabilité qu'un effet néfaste survienne chez un être humain à cause de l'exposition à une substance chimique qui présente un danger¹.

L'évaluation des risques fait face à un nombre de défis toujours croissant, la population étant exposée quotidiennement à des centaines de substances chimiques, par des sources de contamination et des voies d'exposition multiples, et chez certains avec une sensibilité accrue (par exemple aux âges extrêmes de la vie). Ces substances sont parfois retrouvées en mélange et leur coexposition favorise des risques cumulatifs et des interactions toujours plus complexes. Ces problématiques nécessitent de développer des outils et des méthodes pour agir de façon adaptée² : le risque résultant de la combinaison de l'exposition et du danger, ces deux informations doivent être prises en compte pour chacune des

substances, afin d'identifier quels sont les composés et les effets responsables du risque dans sa globalité².

La toxicologie s'intéresse à la réponse des organismes vivants exposés à des agents toxiques grâce à l'étude du devenir des substances toxiques et de leurs effets sur l'organisme³. En définissant la biodisponibilité de ces substances par les diverses voies d'entrée dans l'organisme, la toxicologie permet de comprendre comment une dose absorbée se distribue dans les organes (études de toxicocinétique) et quels sont les effets associés (études de toxicodynamie). Il faut pour cela connaître les mécanismes d'actions associés à ces substances, l'interaction potentielle entre plusieurs substances et les biomarqueurs d'effet et d'exposition³. L'identification et la caractérisation des dangers permettent dans de nombreux cas de proposer une dose seuil à partir de laquelle les effets toxiques apparaissent, mais l'obtention de cette dose peut être limitée par des éléments tels qu'une courbe effet dose non monotone ou des mélanges des substances (effets « cocktails »)¹.

Le risque toxicologique associé aux expositions environnementales fait partie intégrante des préoccupations pour la santé humaine. La contamination de la chaîne alimentaire, de l'air, de l'eau ou des sols, la survenue d'un accident industriel majeur et ses effets à long terme occupent l'attention du grand public à travers les médias³. Ces problématiques de santé publique sont aujourd'hui au cœur des priorités législatives des gouvernements, notamment européens.

La toxicologie fournit des données qui relient les expositions et la survenue d'effets toxiques (la relation dose-réponse), et l'évaluation des risques en matière de santé va utiliser ces données pour réaliser des prédictions et aider à la prise de décision vis-à-vis des dommages en santé publique⁴. Cette discipline utilise les données qui relient les expositions et la survenue d'effets toxiques (la relation dose-réponse), fournies par la toxicologie, pour réaliser des prédictions.

L'étude du risque toxique peut être abordée selon diverses approches, avec chacune leurs données distinctes, et les avantages de l'une compensent les limitations d'une autre. Ainsi, l'association de résultats variés fournit des informations complémentaires, qui permettent de caractériser les effets toxiques. La toxicologie expérimentale utilise différents modèles d'étude, qui permettent d'évaluer les risques dans des conditions d'exposition spécifiques et connues, permettant ainsi la caractérisation de la relation dose-effet (la sévérité de l'effet augmente avec la dose, pouvant concerner différents organes) et de la relation dose-réponse (incidence de l'effet à l'échelle d'un groupe ou d'un organisme)⁵.

Historiquement, l'évaluation toxicologique de nouvelles substances chimiques se faisait essentiellement à travers des expérimentations *in vivo* afin de déterminer avant une mise sur le marché s'il existait un risque pour la santé humaine. Les modèles animaux, de par leurs conditions d'expérimentation contrôlées, la faible variabilité interindividuelle et l'anticipation des effets possibles représentent un modèle d'étude pertinent pour l'évaluation du potentiel dangereux d'une substance⁴. Pourtant, les différences d'espèces en terme de toxicocinétique et de toxicodynamie, le coût important engendré par ces études, les problématiques d'éthique animale et le manque de représentativité d'expositions humaines constituent des inconvénients majeurs pour ces modèles, qui doivent donc être inscrits dans une association d'approches complémentaires pour l'identification des dangers⁴. La suite de ce travail se concentrera sur l'état des lieux des méthodes dites « traditionnelles » utilisée en toxicologie expérimentale et leurs alternatives, prometteuses dans un contexte évolutif d'évaluation des risques.

I. Evolution des méthodes traditionnelles d'expérimentation en toxicologie vers les méthodes alternatives

A. État de l'art de l'expérimentation animale en évaluation des risques

Bien que l'observation d'animaux morts ou vivants ait été une pratique courante au long de l'histoire, l'expérimentation animale a été théorisée par Claude Bernard en 1865 dans son ouvrage « Introduction à l'étude de la médecine expérimentale » comme un outil indispensable pour acquérir des connaissances en santé humaine. Selon l'auteur, le concept d'expérimentation sur les animaux remonte en fait à l'Antiquité : le médecin Gallien « institua ses expériences en particulier sur des singes, ou sur de jeunes porcs et il décrit les instruments et les procédés employés pour l'expérimentation »⁶. Dans le chapitre « De la vivisection », il justifie la vivisection d'animaux vivants comme un complément nécessaire à la dissection de cadavres.

Les modèles *in vivo* doivent pouvoir reproduire fidèlement la physiologie humaine, pour que les observations puissent être extrapolées, du moins en partie. L'origine commune de tous animaux rend possible cette fidélité : le métabolisme, ainsi qu'une partie importante du patrimoine génétique, sont conservés au cours de l'évolution. L'expérience démontre la pertinence et la similitude d'un modèle animal avec l'être humain, mais chaque modèle peut présenter des caractéristiques spécifiques. Aucun modèle seul n'est représentatif de l'organisme humain ou capable de répondre à un ensemble complet d'hypothèses. Ainsi, il est nécessaire de combiner les résultats de plusieurs modèles et les adapter au fil du temps⁷.

Les souris, les poissons, les rats et les oiseaux, représentent en 2024, par ordre d'importance en termes d'utilisation, 93 % des animaux utilisés en

laboratoire⁸. Ainsi, les souris représentent l'espèce la plus utilisée en laboratoire : malgré un génome 14 % plus petit que celui de l'humain, 80 % de leur 30 000 gènes codant pour des protéines sont présents dans le génome humain⁹. De plus, leurs différences avec l'humain peuvent être un atout en laboratoire, que ce soit leur taille, la détoxification intense au niveau de leur métabolisme, leur grande capacité reproductive, leur facilité d'élevage en captivité et la possibilité de consanguinité¹⁰.

Plus récemment, le poisson zèbre (*zebrafish*) prend une place centrale parmi les modèles utilisés en toxicologie, notamment grâce à son important génome (26 000 gènes), dont 70 % présente une similitude avec l'humain. Sa petite taille, sa bonne capacité reproductive, sa rapidité de croissance et son développement uniforme rendent son élevage rapide et peu coûteux en laboratoire. De plus, l'organisation et le fonctionnement de son système nerveux est proche de celui de l'humain. Enfin, les larves de *zebrafish* ne sont soumises à la législation et aux préoccupations éthiques de l'utilisation des animaux en laboratoire qu'à partir de leur 5^e ou 6^e jour de vie, moment où la larve commence à se nourrir¹¹.

Ainsi, durant cette fenêtre de quelques jours, cette espèce est considérée comme un modèle alternatif : elle est largement utilisée aujourd'hui pour des activités de criblage à haut débit pour évaluer la toxicité des substances chimiques¹².

Selon l'enquête de 2022 sur l'utilisation d'animaux à des fins scientifiques dans les établissements français, réalisée par le Ministère de l'Enseignement Supérieur et de la Recherche¹³, 39 % des animaux de laboratoire sont utilisés par la recherche fondamentale, 24 % par les études toxicologiques et réglementaires (mise au point, production ou essais d'innocuité des substances chimiques), et enfin 23 % par la recherche translationnelle (interface entre la recherche fondamentale et la recherche clinique). Les derniers 14 % sont répartis entre la maintenance de colonies d'animaux génétiquement modifiés (10,6 %), la

recherche à des fins de conservation des espèces (1,7 %) et l'enseignement (1,7 %) (figure 1).

En toxicologie réglementaire, 64 % des animaux sont utilisés pour la législation encadrant les médicaments à usage humain, 16 % pour les médicaments à usage vétérinaire et 8 % pour les produits chimiques⁸. Enfin, depuis 2009, il est interdit d'utiliser des animaux pour les tests de produits finis cosmétiques et les ingrédients qui les composent¹⁴.

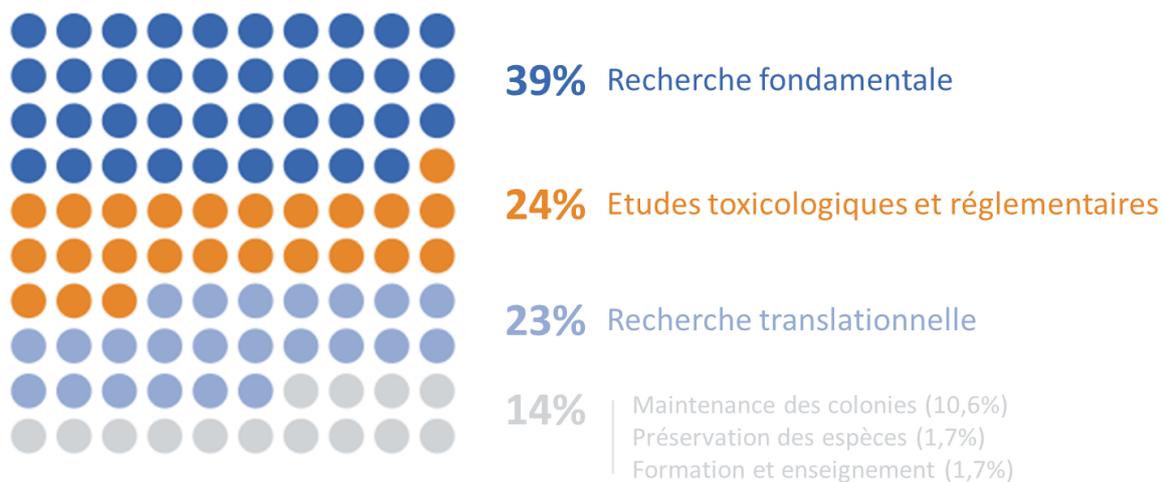


Figure 1: Bénéficiaires des expérimentations animales en France (adapté de l'Enquête statistique sur l'utilisation des animaux à des fins scientifiques, Ministère de l'Enseignement supérieur et de la recherche (2022))¹³

La toxicologie servant à anticiper les risques en termes de santé humaine, elle se doit d'agir en amont, en utilisant différents modèles en laboratoire pour comprendre les phénomènes physiopathologiques engendrés par les toxiques. Le modèle animal permet de réaliser des études complexes irréalisables sur l'humain ; néanmoins, l'utilisation d'organismes modèles n'est pas sans limites, il n'existe pas de substitut parfait et les extrapolations d'un organisme à l'autre doivent être faites avec précaution.

B. Les défaillances de l'expérimentation animale traditionnelle

1. Principales difficultés

L'expérimentation animale demeure une composante indispensable à l'évaluation de la sécurité des agents chimiques, que ce soit pour la nature des effets toxiques, répondant à des conditions rigoureusement contrôlées, la dose responsable des effets toxiques et la pertinence humaine des essais réalisés sur les animaux¹⁵. Cependant, cette approche traditionnelle est confrontée à de nombreuses limites : la variabilité inter-espèces, l'histologie et la physiologie parfois différentes de la physiologie humaine, les pathologies spontanées fréquentes ainsi que les enjeux éthiques limitent le nombre d'animaux utilisés. Les retraits de substances chimiques, ou l'élimination de molécules candidates (« faux négatifs ») pendant la phase préclinique ou clinique¹⁶ sont nombreux. Considérant ces problématiques, il apparaît nécessaire de changer de paradigme quant à l'expérimentation traditionnelle.

Tout d'abord, pour une grande partie des substances chimiques mis sur le marché européen, les données nécessaires à une évaluation des risques fiable et pertinente manquent ou sont incomplètes. En ce qui concerne les substances chimiques produites aux tonnages les plus importants, soumises à des exigences plus strictes, ces lacunes au niveau des données sont souvent dues à un manque de conformité de la part des déclarants, une compréhension limitée des exigences ou encore des difficultés dans le partage des données entre industriels déclarant la même substance¹⁷. De plus, quand les données sont disponibles, elles sont à risque de manquer de fiabilité : une analyse a identifié des problèmes de reproductibilité des huit tests réglementaires les plus courants, présents dans les lignes directrices de l'OCDE, pour la toxicité aiguë et dermique¹⁸. En ce qui concerne les substances chimiques commercialisées au tonnage le plus faible (1 à 10 tonnes par an selon le règlement européen REACH¹⁹), mais qui peuvent

néanmoins représenter un tonnage commercialisé très élevé, les informations réglementaires nécessaires n'incluent qu'un ensemble d'indicateurs basiques afin de juger si les utilisations pour lesquelles une substance est enregistrée engendrent un risque potentiel pour la santé humaine²⁰.

En excluant les interrogations éthiques que posent les tests sur les animaux, ainsi que la pertinence scientifique de la réalisation de ces tests, effectuer des tests sur les animaux représente un coût important et un temps considérable compte tenu des très nombreuses substances à tester, déjà sur le marché ou attendues dans les années à venir.

Pour les nouvelles substances et produits chimiques développés, les ressources et le temps nécessaires à la réalisation de tests animaux ne sont souvent pas en accord avec les exigences de l'industrie²⁰. Il est fondamental d'améliorer l'efficacité de la collecte de données toxicologiques, l'évaluation des risques des substances chimiques déjà sur le marché, et cela ne peut reposer uniquement sur l'expérimentation animale²⁰.

La correspondance entre modèles animaux et maladies humaines représente également un obstacle à la fiabilité de la transposition. En effet, au cours de ces expérimentations, les pathologies humaines sont artificiellement induites chez les animaux : la difficulté est donc de reproduire toute leur complexité. Ainsi, même si l'expérimentation animale est bien décrite et standardisée, la transposition des résultats de l'animal vers l'être humain peut s'avérer source de biais.

La question de la pertinence et de la prédictibilité du modèle animal se pose aussi lors d'essais cliniques : en effet, il est estimé que dans près de 9 cas sur 10, les candidats médicaments testés sur les animaux en phase précliniques échouent lors des phases cliniques sur l'humain²¹. Même si la toxicité observée chez l'animal est très proche du résultat sur l'humain, les différences entre ces espèces en matière de comportement, de physiologie, de génétique ou pharmacocinétique

limite la fiabilité des tests animaux. Les prédictions réalisées sur des rongeurs ne sont qu'à 43 % représentatives d'un organisme humain, et à moins de 30 % quand il s'agit de prédire des effets indésirables de médicaments²². Le taux d'échec en essai clinique est très important, notamment dû au manque de prédictivité des modèles animaux : selon le « National Institute of Health » (NIH), aux Etats-Unis, 30 % des substances qui passent les tests animaux sont toxiques chez l'humain et 60 % des molécules efficaces sur les animaux ne le sont pas chez les humains²¹. Le modèle animal est peu prédictif des effets chez l'humain²³ et il est possible d'observer des différences dans les effets d'un produit chimique entre deux espèces animales étudiées : il y a, notamment, pour les tests de toxicité sur la reproduction, seulement 60 % de cohérence entre espèces²⁴. Outre ces difficultés de cohérence, des divergences au sein même d'une espèce ont été démontrées : en effet, une étude a révélé que des rats issus de la même souche mais provenant d'un fournisseur différent présentent des altérations variées suite à une lésion de la moelle épinière²⁵. Ces différences influencent donc la réponse de l'animal à une atteinte physique et par extension les résultats à un traitement. De plus, certains auteurs soutiennent l'existence d'une « crise de la reproductibilité »^{26,27} : les critères méthodologiques nécessaires pour contrôler les biais dans les essais cliniques ne sont pas toujours respectés, ce qui entraîne fréquemment une faible puissance statistique, un manque de prise en compte des facteurs environnementaux, et parfois la non-publication de résultats négatifs.

L'utilisation d'animaux pour l'expérimentation reste donc un pilier essentiel dans la recherche fondamentale, l'évaluation des risques et le contrôle de la qualité des produits de santé. Pourtant, la confiance en ce modèle traditionnel s'épuise face aux contraintes en matière de qualité précédemment évoquée, mais aussi par l'évolution constante des perceptions culturelles qui se caractérise, depuis le début du XIX^e siècle, par une sensibilisation croissante de la société au bien-être animal et au respect des droits des animaux. Dans ce contexte, la

réglementation concernant les pratiques liées à l'expérimentation animale repose sur deux objectifs principaux : limiter autant que possible l'utilisation d'animaux et, lorsque leur utilisation est inévitable, déterminer le seuil acceptable de souffrance animale. Cette démarche constitue la base de la démarche éthique, aujourd'hui fondamentale à l'expérimentation animale.

2. Limites éthiques

La législation n'a cessé d'évoluer concernant l'éthique animale depuis le XX^e siècle. En effet, l'unique réglementation entourant l'expérimentation animale datait de 1850, avec la loi Gramont, qui imposait une amende à ceux qui ne respectaient pas le bien-être animal, sans pour autant détailler les conditions d'expérimentation²⁸. En 1959, William Russell, zoologiste, et Rex Burch, microbiologiste, publiaient « The Principles of Humane Experimental Technique »²⁹, qui pose les fondements de la prise en compte de l'éthique animale dans les expériences scientifiques. Ces principes sont connus sous le nom des 3R : Remplacer, Réduire, Raffiner. Ils visent à minimiser l'utilisation d'animaux, en remplaçant dès que possible les animaux par d'autres méthodes, en réduisant la quantité d'animaux et en raffinant les méthodologies déjà existantes.

Le premier terme introduit par Russell et Burch, « Remplacer », représente la nécessité d'utiliser des méthodes alternatives à la place des modèles vivants. Il est donc nécessaire de s'appuyer sur des cellules (*in vitro*), sur des modèles informatiques (*in silico*), sur des réactions chimiques (*in chemico*), ou sur des prélèvements de tissus (*ex vivo*)³⁰.

Malgré l'absence d'un remplacement complet des espèces animales, le terme « Réduction » vise à la réduction des répétitions inutiles d'études antérieures, à la rédaction d'un protocole expérimental avant toute expérimentation pour éviter d'autres essais inutiles³⁰ et à la limitation de l'utilisation d'animaux aux

expériences essentielles. Ces dernières font l'objet de débats quant à leur définition, car s'il est essentiel d'utiliser le plus petit nombre possible d'animaux, il faut néanmoins éviter que cela nuise à l'obtention de résultats significatifs : une analyse de « puissance statistique » doit permettre de définir le déroulé d'une expérience conduisant à des résultats exploitables³¹.

Enfin, « Raffiner » fait référence à l'ensemble des méthodes qui visent à réduire, soulager ou supprimer la souffrance et la détresse des animaux, et qui visent à améliorer leur bien-être en général. Il est nécessaire de définir un ensemble de bonnes pratiques qui permettent de minimiser les dommages sur les animaux³¹. Par exemple, cela conduit à choisir soigneusement le modèle à utiliser en amont de l'expérience, améliorer les conditions d'élevage et d'hébergement, mettre en place un protocole de réduction du stress et établir des « points limites » à la procédure³⁰. Ces derniers font référence au moment où la souffrance d'un animal doit être arrêtée ou minimisée : ce point limite doit donc pouvoir être reconnu par les expérimentateurs et fixé le plus précocement possible³². Raffiner au cours de l'expérimentation peut se traduire par des mesures non invasives comme de l'imagerie, des soins adéquats tout au long de la procédure, l'analgésie, la réduction du temps de certaines études et des procédures d'euthanasie appropriées³⁰. De plus, des résultats exploitables dépendent directement du bien-être animal : l'adage « happy animals make good science », énoncé par Trevor Poole en 1997, résume l'importance de l'éthique dans les bonnes pratiques scientifiques³³.

La loi de 1986, qui décrétait une sanction pénale en cas de maltraitance lors des expérimentations, a été remplacée par la directive 2010/63/UE³⁴ relative à la protection des animaux utilisés à des fins scientifiques. Elle insiste sur le fait que « les pratiques doivent être limitées aux expérimentations strictement nécessaires et prendre en compte le fait que les animaux sont des êtres sensibles, sujets à la

douleur et ayant des besoins physiologiques et comportementaux propres à chaque espèce », avec un objectif de remplacement total des procédures actuelles³⁵. Au 1^{er} janvier 2013, cette directive a été complétée par la publication d'une ordonnance dans le Code rural qui soumet à une autorisation préalable, suite à une évaluation par une autorité compétente, tout projet qui implique l'utilisation invasive ou non d'un animal à des fins expérimentales²⁸. Dans un contexte de sensibilité croissante au bien-être animal, les méthodes alternatives à l'expérimentation *in vivo* permettent d'apporter des réponses aux attentes sociétales.

3. Développement nécessaire des méthodes alternatives

Le rapport de la Commission Européenne réalisé en 2019 sur l'utilisation des animaux à des fins scientifiques représente un premier bilan des efforts menés par les différentes politiques gouvernementales depuis la directive 2010/63/EU³⁶. Analysant les chiffres collectés par les laboratoires entre 2015 et 2017, ce rapport conclut que, malgré une politique européenne rigoureuse en matière de protection des animaux, ces derniers continuent d'être utilisés systématiquement dans les essais concernant les substances chimiques (plus de 230 000 tests sur animaux en 2017 pour satisfaire la législation)³⁶. Partant de ce constat, dans sa stratégie pour le développement durable des produits chimiques, la Commission Européenne s'est engagée dans le développement de nouveaux matériaux, de modèles et de méthodes pour l'évaluation des produits chimiques, dans une volonté d'abandonner les études sur les animaux³⁷.

Suite aux nombreuses questions posées par l'évolution de la recherche en toxicologie, il est essentiel de développer de nouvelles manières d'expérimenter et, tant dans l'industrie pharmaceutique et chimique que dans le processus de prise de décision réglementaire, de réaliser une évaluation des risques plus fiable,

plus rapide, et moins coûteuse et qui minimise le recours à l'expérimentation animale. Outre ces exigences éthiques, il s'agit également d'une nécessité réglementaire dans un contexte législatif qui préconise un changement dans la manière de réaliser les tests toxicologiques.

Ainsi, des tests minimisant l'impact animal doivent être introduits afin de participer à l'évaluation du danger engendré par les substances chimiques. Le terme « méthodes alternatives » a été introduit pour la première fois en 1978 par David Smyth, physiologiste anglais, et regroupe des suggestions pratiques et constructives permettant de satisfaire aux principes des « 3R »³⁸. Ainsi, ces méthodes poseront les fondements de ce qui est aujourd'hui appelé en toxicologie prédictive les « nouvelles approches méthodologiques », ou NAM, qui continuent à utiliser des animaux mais en s'inscrivant dans une démarche de réduction du recours aux expérimentations *in vivo*⁸. Elles incluent notamment les modélisations mathématiques (*in silico*) et les tests sur les cellules (*in vitro*).

Les données fournies par les méthodes alternatives permettent d'approfondir la compréhension de l'origine de la toxicité d'une substance chimique, parfois même de fournir des données plus pertinentes que les tests traditionnels en termes de toxicité humaine ou environnementale³⁹.

Au-delà des limites éthiques que pose l'expérimentation animale, les attentes sociétales importantes concernant la santé et l'environnement (notamment les effets des pesticides et des perturbateurs endocriniens) ainsi que les progrès scientifiques dans la compréhension des mécanismes d'effets (effets à faible dose, effets non linéaires et effets « cocktails ») nécessitent une réévaluation des approches traditionnelles en toxicologie¹⁵. Le nombre croissant de substances chimiques, les nouveaux matériaux disponibles (comme les nanomatériaux), de nouveaux sujets de préoccupations (neurotoxicité développementale, immunotoxicité, perturbation endocrinienne...) questionnent le système actuel⁴⁰.

La pertinence des modèles rongeurs, perçus comme le « gold standard » des expérimentations, notamment de toxicité répétée et de cancérogénicité, est également remis en question^{41,42}. Les différences inter espèces, au sein même de la famille des rongeurs, l'extrapolation peu fiable des animaux de laboratoire à l'être humain, le manque de clarté dans les résultats *in vivo* et le manque de reproductibilité mettent en doute la pertinence de ces méthodes d'essai pour l'évaluation des risques en santé humaine^{18,43}.

Les « nouvelles approches méthodologiques », qui intègrent tout ce qui ne relève pas des approches traditionnelles dans l'expérimentation, semblent pertinentes pour répondre à ces problématiques. Leur implication grandissante dans l'élaboration de directives d'essai mises en place par les autorités européennes laisse entrevoir une standardisation susceptible de modifier profondément la réglementation des substances chimiques⁴⁴.

C. Émergence des « Nouvelles Approches Méthodologiques » (NAM) en toxicologie

1. Définitions et principes

Afin de répondre aux attentes éthiques et réglementaires en expérimentation animale, il est nécessaire de disposer de méthodes reconnues et validées. Les NAM, nouvelles approches méthodologiques (en anglais « New Approach Methodologies »), représentent en toxicologie un défi pour la recherche et le développement scientifiques, ainsi que l'évaluation de la sécurité des produits chimiques pour la santé humaine et l'environnement. Elles permettent en effet d'améliorer les connaissances sur l'interaction entre les produits chimiques et les systèmes biologiques exposés à ces produits chimiques⁴⁵. Il n'existe actuellement aucune définition harmonisée des NAM, au niveau européen ou mondial. Selon l'Agence européenne des produits chimiques (ECHA), qui garantit la sécurité des

produits chimiques en Europe, les NAM représentent « toute technologie, méthodologie, approche ou combinaison pouvant fournir des informations sur les dangers chimiques et l'évaluation des risques sans l'utilisation d'animaux, comprenant les approches *in silico*, *in chemico*, *in vitro* et *ex vivo* »⁴⁶. Les NAM prennent en compte les informations relatives à l'exposition et au mécanisme d'action, les propriétés physicochimiques et les propriétés d'absorption, de distribution, de métabolisme et d'élimination (ADME) des substances à évaluer afin d'étayer une conclusion sur une propriété par une approche fondée sur le poids de la preuve (ou « weight-of-evidence »)⁴⁷ (figure 2). Cette dernière approche permet de formuler une synthèse des preuves disponibles, parfois de qualités hétérogènes, dans le but de déterminer la plausibilité d'hypothèses formulées sur les données⁴⁸. La lecture croisée, plus communément appelée en Europe « read-across », permet de collecter et de croiser des données pertinentes de substances analogues afin de calculer une prédiction de toxicité⁴⁹.



Figure 2: Schéma général des composantes des NAM.

La toxicologie s'oriente désormais vers la compréhension approfondie des voies de toxicité qui mènent à des effets néfastes. Mieux comprendre les mécanismes biologiques et leur progression vers la manifestation d'une toxicité

permet d'anticiper les perturbations éventuelles causées par l'exposition à des substances chimiques.

Les « chemins de l'effet néfaste » (« Adverse Outcome pathways » ou AOP) sont des outils développés pour caractériser un danger, non plus en l'identifiant par des tests chez des animaux exposés à des doses élevées mais en étudiant les voies biologiques grâce à des modèles NAM⁵⁰. Ils proposent une approche pour analyser le mécanisme conduisant au risque toxicologique : un enchaînement de divers événements clés (« Key events ») conduisent à la survenue d'un effet néfaste⁵¹. Étant donné la complexité des systèmes biologiques, les AOP nécessitent une expertise scientifique multidisciplinaire, reposant sur une analyse pertinente des données. Les AOP ne décrivant pas de façon exhaustive les réactions et les phénomènes cellulaires présents au sein d'un être vivant, la caractérisation de ces phénomènes biologiques doit s'appuyer sur des méthodes NAM, relativement simples et accessibles⁵². Ils décrivent également les interactions moléculaires au niveau du site d'action ainsi que la séquence de réactions se produisant au niveau des cellules, des tissus, de l'organisme et de la population⁵⁰. Les AOP permettent de décrire une séquence qui commence par un événement moléculaire initiateur, au cours duquel une substance interagit avec une cible biologique, suivi d'une série d'événements clés se produisant à différents niveaux de l'organisme et résultant en un effet néfaste (figure 3)⁵⁰. Ils constituent ainsi une analyse critique et ciblée des résultats obtenus grâce aux méthodes alternatives.

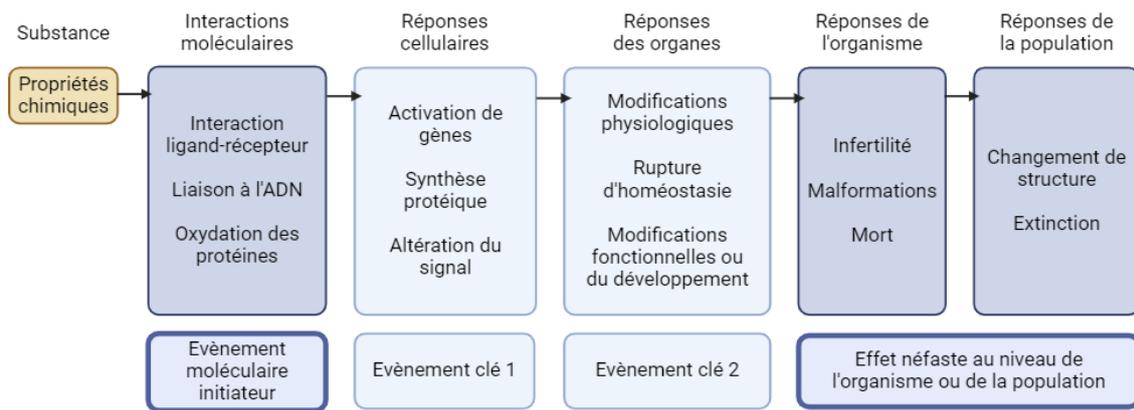


Figure 3: Schéma conceptuel des principales caractéristiques d'un AOP, selon Ankley (2010)⁵³. Chaque AOP commence avec un évènement moléculaire initiateur (par exemple l'interaction entre un ligand et un récepteur), suivi de divers évènements clés à plusieurs échelles de l'organisme, pour aboutir à un effet néfaste (par exemple un effet sur le développement ou la reproduction). Les trois premières étapes représentent la « voie de toxicité » : la réponse cellulaire n'étant pas adaptée à la perturbation, elle induit une toxicité⁵³.

Les NAM peuvent également inclure des expérimentations animales, mais qui permettront une prise de décision plus rapide en matière de sécurité et nécessiteront moins de ressources (moins chères, moins de temps pour les analyses et moins d'animaux utilisés)⁴⁴. Les avancées en matière d'imagerie médicale (IRM, scintigraphie, tomographie, ultrasons, etc...) et de télémétrie (enregistrement en continu de paramètres biologiques) favorisent notamment l'amélioration des techniques *in vivo*, en permettant par exemple de choisir des espèces moins sensibles à la douleur¹⁵.

Ainsi, ces méthodes respectent le principe des 3R - remplacer, réduire et raffiner l'expérimentation animale - minimisant ainsi l'utilisation des animaux, aidant à améliorer la compréhension mécanistique des effets toxicologiques des produits chimiques sur les systèmes biologiques et favorisant l'extrapolation à l'être humain des résultats obtenus dans les études animales⁵⁴. Leur pertinence doit être démontrée ; elle doit être similaire ou supérieure à celle des approches existantes : les résultats doivent être reproductibles, le test pertinent et l'approche adaptée à l'objectif (« fit for purpose »)⁴⁴.

Les NAM ne sont pas de nouvelles méthodes, mais leur utilisation dans un contexte réglementaire visant à la diminution de l'expérimentation animale traditionnelle pour l'évaluation des dangers des substances chimiques est relativement récente. Leur importance scientifique pour l'évaluation des dangers et des risques étant reconnue par les agences de réglementation depuis le début des années 2000, la confiance portée dans l'utilisation des NAM pour la prise de décision réglementaire ne cesse d'augmenter^{55,56}. Dans ce contexte, il est essentiel de développer des outils au niveau cellulaire et informatique robustes pour une identification fiable et précise des dangers dans le cadre de l'évaluation réglementaire des produits chimiques et pharmaceutiques.

2. Types de NAM

Afin de remédier aux différentes lacunes précédemment mentionnées, de plus en plus de NAM ont été mises au point ces dernières années, avec pour objectif de fournir un système d'évaluation du danger qui soit scientifiquement meilleur, plus efficace et éthique. De nombreuses méthodes alternatives à l'expérimentation animale ont été développées depuis le début du 21^e siècle, avec une meilleure identification des dangers et la compréhension mécanistique des effets toxicologiques.

En général, les NAM peuvent être classées en différents groupes de méthodes, en fonction des supports utilisés.

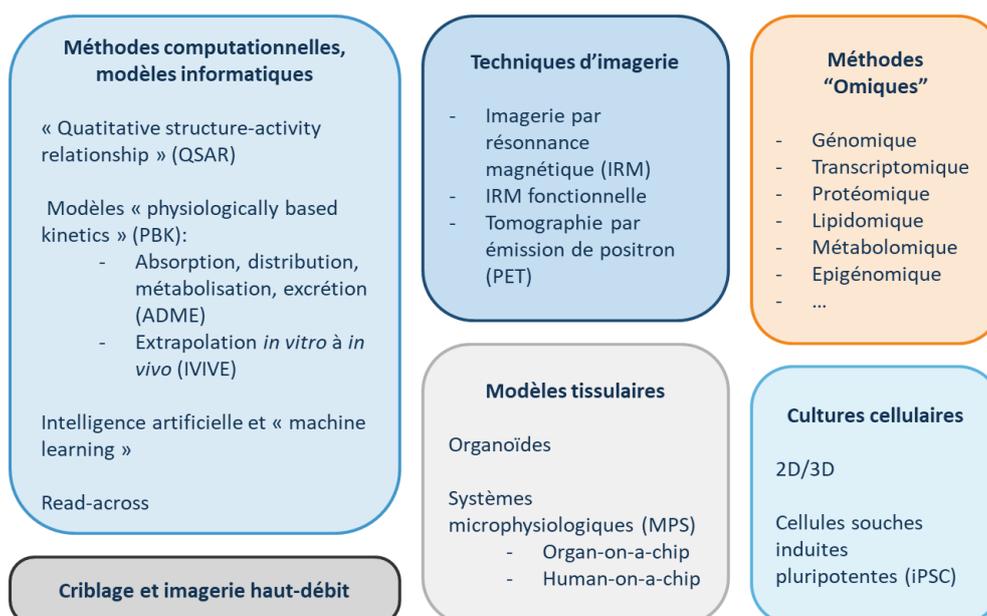


Figure 4: Liste non exhaustive de méthodes, outils et applications NAM (adapté de Schmeisser et al. (2023)²⁰).

De tels modèles NAM incluent les méthodes de criblage et d'imagerie haut-débit, le champ des applications « omiques » (grands volumes de données fournissant des informations sur les processus biologiques à différentes échelles de la cellule⁵⁷), les cultures cellulaires, les organoïdes et autres systèmes microphysiologiques, les modèles de machine learning et d'intelligence artificielle, les prédictions QSAR et les « read-across » (figure 4).

Les techniques *in vitro* consistent en l'utilisation de cellules en culture, avec des modèles très diversifiés⁷. Parmi eux, les modèles 2D sont soit des cultures primaires (préparations directement extraites d'organes) soit des lignées (transformées et/ou immortalisées pour avoir une homogénéité de réponse et des capacités reproductives importantes)⁷. Ils constituent les modèles les plus utilisés dans la recherche, mais présentent comme limite de ne pas prendre en compte l'environnement cellulaire dans lequel les cellules évoluent au sein des tissus⁵⁸. Les modèles 3D permettent de s'affranchir de cette limite, se développant sur une matrice qui permet aux cellules d'interagir avec leur environnement dans les trois dimensions, se rapprochant des conditions réelles. Certains types cellulaires

possèdent des propriétés de différenciation intéressantes comme les cellules Caco2, cellules épithéliales de la barrière intestinale humaine dérivées d'un adénocarcinome colique, utilisées pour étudier le passage des toxiques au niveau intestinal⁵⁹. Les données générées *in vitro* qui utilisent des cellules humaines ou des organoïdes, peuvent directement être utilisées pour caractériser les cibles cellulaires des substances chimiques et les mécanismes moléculaires impliqués dans leur toxicité. Les cibles et les mécanismes moléculaires sont décrits dans les AOP^{52,60}.

La toxicologie *in silico* (ou toxicologie computationnelle) utilise des méthodes informatiques pour organiser, simuler, analyser, modéliser, visualiser ou prédire la toxicité des substances chimiques⁶¹. De façon non exhaustive, elle inclut :

- Des bases de données de substances chimiques,
- Des logiciels pouvant générer des descripteurs moléculaires (variables numériques, descriptives de la structure moléculaire),
- Des outils de simulation des dynamiques moléculaires et des systèmes biologiques,
- Des logiciels pour générer des modèles de prédiction,
- Des applications générant des prédictions de toxicité,
- Des outils de visualisation des données⁶².

Les bases de données comprenant des résultats d'anciennes expérimentations *in vivo* ou *in vitro* étant nécessaires afin d'obtenir des résultats concernant les substances étudiées, les domaines *in vivo*, *in vitro*, et *in silico* sont liés et interagissent entre eux⁶³. Parmi les techniques de toxicologie *in silico*, les modèles QSAR (Quantitative Structure Activity Relationship) sont les plus courants à l'heure actuelle ; ils seront utilisés dans la suite de ce travail. La modélisation QSAR décrit la relation entre les caractéristiques structurales d'une substance chimique et son activité toxicologique⁶⁴. Les modèles PBK (« Physiologically Based Kinetics ») sont

également des outils privilégiés en techniques *in silico*, ils permettent de décrire l'absorption, la distribution, la métabolisation et l'élimination (propriétés ADME) de substances chimiques¹⁵. Ils peuvent également se baser sur les modèles d'extrapolation *in vitro-in vivo* (IVIVE), qui utilisent des résultats obtenus *in vitro* afin de prédire des résultats *in vivo* ⁶⁵.

La toxicologie *in chemico* utilise les résultats de réactions chimiques pour fournir des informations sur la réactivité de la molécule étudiée et ainsi pouvoir prédire sa toxicité. Afin de faciliter l'identification des dangers, ces informations obtenues *in chemico* sont souvent traduites en des outils *in silico* ⁶⁶. L'évaluation de la sensibilité cutanée est par exemple largement décrite par l'OCDE, qui a mis en place trois lignes directrices spécifiques (442C ⁶⁷, 442D ⁶⁸ et 442E ⁶⁹), intégrant des méthodes de tests *in chemico*. A chaque ligne directrice correspond un AOP, dont l'évènement initiateur du danger peut être décrit par un test NAM²⁰. Le test DPRA (Direct Peptide Reactivity Assay) est un exemple de test *in chemico* intégré dans les AOP sur la sensibilité cutanée. Il vise à mesurer la concentration restante en cystéine après incubation avec une molécule pendant 24 h pour évaluer la sensibilité cutanée à une substance chimique, selon la ligne directrice 442C de l'OCDE⁶⁷. Ainsi, les méthodes alternatives ne sont pas limitées à un seul type de test ; l'intégration de plusieurs NAM au sein d'un critère d'évaluation, et donc d'un AOP, permet de surmonter les limites de chaque méthode.

Les cultures *ex vivo* représentent par exemple un modèle où des tranches de tissus entières sont cultivées, ce qui permet de conserver en partie leur organisation cellulaire et ainsi se rapprocher du processus métabolique de l'*in vivo*⁷⁰. Cet isolement de cellules adultes humaines, considéré comme faisant partie des méthodes *in vitro*, permet de s'affranchir d'une étude *in situ* tout en étudiant leur développement en temps réel. Les méthodes *ex vivo* facilitent notamment l'étude des cellules souches par la visualisation des facteurs de croissance et leur

caractéristiques⁷¹. Certains modèles *ex vivo* comme des tranches d'œil de bovins sont également utilisés pour évaluer des produits corrosifs, selon la ligne directrice 437 de l'OCDE⁷².

D'autres technologies, telles que les méthodes « omiques », font également partie des disciplines émergentes en toxicologie. Cette terminologie s'utilise pour des analyses qui vont avoir lieu à différents niveaux de la cellule : fondée sur l'idée qu'un évènement toxicologique va entraîner la modification de l'activité protéique et de la fonction cellulaire, cette méthode permet d'évaluer les profils de modification obtenus après exposition à une substance chimique⁷³. Ces données « omiques » sont collectées à travers différentes disciplines de la toxicologie afin de caractériser le comportement des gènes (transcriptomique), des protéines (protéomique), des métabolites (métabolomique) et autres composants des voies biologique⁷⁴. La collecte de données « omiques » lors d'un test de toxicité à court terme permet de prédire les effets à long terme. En comprenant mieux les voies entraînant une perturbation, il est possible de choisir les substances chimiques qui méritent plus ample évaluation, ce qui facilite la prise de décision, réduisant ainsi le nombre d'animaux et la durée de l'expérimentation⁷⁵.

Le futur des expérimentations en toxicologie repose surtout sur l'arrivée des modèles *in silico*, qui permettent d'utiliser des données déjà existantes et de les mettre en lien afin de prédire le risque toxique associé à l'exposition à de nombreuses substances. Les méthodes *in vitro* et *ex vivo* offrent également des pistes d'exploration en tant qu'alternatives pour un remplacement total du modèle animal, ce sont les seules qui permettent de travailler avec des modèles humains, s'affranchissant ainsi théoriquement des variabilités inter espèces observées habituellement⁷⁶.

3. Avantages et inconvénients des différentes approches

En ce qui concerne l'efficacité, la capacité prédictive, mais aussi la sensibilité pour détecter les effets à l'échelle humaine, les tests *in vivo* actuellement disponibles sont souvent remis en question ⁷⁷. Plutôt que de remplacer simplement les tests sur les animaux par des tests sans animaux, on estime en effet qu'une intégration de différents tests NAM permettrait d'imiter la complexité biologique et ainsi d'améliorer les prédictions toxicologiques vis-à-vis de la santé humaine⁷⁸.

Les modèles *in vitro* présentent de nombreux avantages, notamment leur rapidité et leur facilité d'installation et de mise en place, leur aspect en général peu onéreux et le grand nombre de modèles existants. La possibilité d'utiliser des cellules humaines, qui peuvent provenir de différents organes et tissus, et la facilité de modélisation des résultats expérimentaux (peu de paramètres biologiques sont à prendre en compte) rend cette méthodologie très intéressante. De plus, les tests *in vitro* nécessitent peu de substance à tester, produisent peu de déchets parfois dangereux à traiter et s'inscrivent dans une démarche éthique ⁴⁰. Pourtant, ne s'agissant pas d'un système entier, ces modèles ne reflètent pas les effets sur l'organisme dans son ensemble, les interactions dans les organes et entre les différents types de cellules, il est ainsi compliqué de modéliser les conditions normales de physiologie humaine⁴⁰. Enfin, les effets pharmacocinétiques ne peuvent pas être identifiés et les contaminations des cultures sont fréquentes⁷.

Les tests *in silico* présentent l'avantage d'être moins coûteux que les méthodes traditionnelles, ne requérant pas d'équipement ou de matériel onéreux. De plus, la recherche *in silico* se fait de façon rapide, les simulations lancées sur ordinateur représentant en général quelques heures de tests, contrairement aux semaines ou mois que nécessitent les tests sur animaux. De plus, elle permet aux experts d'observer des phénomènes impossibles à couvrir en condition réelle

(comportement de particules, températures extrêmes...). De grandes quantités de composés chimiques peuvent être identifiées en même temps, et des prédictions sur les comportements des systèmes biologiques peuvent être réalisées, permettant aux scientifiques de concevoir des expérimentations efficaces. La recherche *in silico* ne pose pas de risques en termes de sécurité des expérimentateurs, elle produit des résultats reproductibles et facilement partagés entre chercheurs. Ainsi, elle permet de compléter la recherche traditionnelle, apportant un point de vue différent, difficile à obtenir en se reposant uniquement sur les méthodes animales⁷⁹.

Malgré ses nombreux avantages, les expériences *in silico* présentent certaines limitations, comme démontrées au cours d'une étude⁸⁰ utilisant simulations et modèles mathématiques. Tout d'abord, les études *in silico* manquent de mise en situation de vie réelle, ils ne prennent pas en compte la complexité et la dynamique de tous les contextes biologiques étudiés. De plus, les tests sont limités à la résolution d'une hypothèse ou d'une prédiction ciblée, plutôt que de couvrir la compréhension d'un système physicochimique ou biologique. Les résultats peuvent manquer de justesse, les modèles informatiques ne pouvant reproduire à l'identique la variabilité des systèmes. Enfin, la reproduction et la validation de résultats peut être difficile à mettre en place, contrairement à l'expérimentation *in vivo* et *in vitro* au cours desquelles les tests peuvent être contrôlés.

D. Cadre réglementaire et intégration des NAM dans l'évaluation des risques

1. Réglementations actuelles

Depuis 2007, le règlement Européen REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) impose aux industriels de fournir des

informations sur les caractéristiques physico-chimiques, toxicologiques et écotoxicologiques des substances chimiques, avant leur mise sur le marché⁸. Ce règlement introduit un changement de mentalité majeur : ce n'est plus aux autorités de démontrer la toxicité d'une substance mais aux industriels de prouver la maîtrise des risques⁸¹. Les autorités peuvent ainsi interdire une substance dont l'innocuité pour la santé humaine et l'environnement n'est pas démontrée, restreindre son utilisation ou la soumettre à une autorisation préalable.

Étant donné le grand nombre de substances à évaluer, et le recours aux animaux qui en découlent, le règlement REACH souligne l'importance de « promouvoir l'innovation et le développement des méthodes alternatives pour l'évaluation des dangers liés aux substances » afin d'évaluer les risques des substances chimiques pour la santé et l'environnement⁸². Bien que les méthodes alternatives ne soient pas obligatoires, le règlement REACH préconise le passage de l'expérimentation *in vivo* à des approches non animales, avec des exigences relatives à la protection des animaux utilisés à des fins scientifiques (directive 2010/63/UE)²⁹.

Il est recommandé d'éviter au maximum l'utilisation d'animaux en favorisant d'autres méthodes qui sont validées ou reconnues. De plus, pour éviter les tests redondants et ainsi diminuer l'utilisation des animaux, le règlement REACH exige le partage des données entre industriels (« one substance, one assessment », décrit dans le « European Green Deal »⁸³). Enfin, les tests sur les vertébrés doivent être spécifiquement demandés et approuvés par l'Agence européenne des produits chimiques (« European Chemicals Agency », ou ECHA). Cette dernière veille à ce que la législation européenne sur ces substances soit bien appliquée, contribue à la validation des méthodes et des approches et encourage le développement des méthodes alternatives dans le cadre de REACH⁸².

Dans sa présentation d'ouverture du congrès annuel de la Société Française de Toxicologie (SFT 2023), Anne Gourmelon, administratrice principale pour les lignes directrices fournies par l'OCDE, évoquait les facteurs clé nécessaires à la mise en œuvre de la feuille de route de l'Union européenne. Il est ainsi nécessaire pour la recherche publique d'utiliser des approches et des méthodes en adéquations avec les réglementations REACH et CLP (classification, étiquetage et emballage), qui demandent l'utilisation de données adéquates, fiables et scientifiquement valides. La disponibilité d'approches alternatives entraîne une possibilité de changement de réglementation, qui elle-même motive le développement de nouvelles méthodes ⁸⁴. Selon l'OCDE, il existe des lignes directrices d'essai pour l'utilisation des NAM, qui peuvent ainsi être standardisées : après validation, elles doivent être couvertes par « l'acceptation mutuelle des données » (un accord entre pays membres pour un partage et une acceptation des données générées selon des réglementations communes). Cette acceptation mutuelle des données permet ainsi d'économiser 309 millions d'euros par an, mais également d'éviter l'utilisation de 10 000 animaux ⁴⁴.

Il existe aujourd'hui approximativement 150 lignes directrices rédigées par l'OCDE, en ce qui concerne les propriétés physico-chimiques, l'écotoxicologie, le devenir des substances dans l'environnement et les effets sur la santé humaine, dont 35 qui comportent des approches non animales ⁴⁴. Le partage des données, la confiance croissante dans les NAM, l'innovation entourant l'évaluation du risque chimique et la diversité des tests existants favorisent les initiatives internationales en matière de réglementation entourant ces nouvelles approches méthodologiques.

2. Initiatives et directives réglementaires en faveur des NAM

Les mandats législatifs actuels de la Commission européenne, de l'Agence européenne des produits chimiques (ECHA) et de l'Autorité européenne de sécurité des aliments (EFSA) exigent de renforcer l'évaluation des substances chimiques sans augmenter l'expérimentation animale. Ces méthodes d'essai alternatives peuvent aider à découvrir les mécanismes sous-jacents aux effets des substances chimiques sur la santé humaine tout en réduisant l'écart entre le nombre de substances utilisées et évaluées. La mise en place de NAM dans les stratégies d'essais réglementaires est nécessaire pour se conformer aux 3R mais aussi limiter les préoccupations associées aux différences inter-espèces⁸⁵. Ainsi, l'utilisation des NAM pour améliorer l'évaluation de la sécurité des substances chimiques est grandissante dans de nombreux secteurs industriels.

Les acteurs industriels de manière générale et ceux impliqués dans la régulation ont l'habitude, du moins en interne, d'utiliser les données générées par des NAM à des fins de criblages préliminaires de molécules et de priorisation de stratégies d'évaluation, notamment grâce au « read-across », permettant de combler les lacunes en matière de données⁸⁶.

Le projet PARC (« Partnership for the Assessment of Risks from Chemicals »), lancé par la Commission Européenne en 2022, a pour volonté de promouvoir la coopération européenne, de faire progresser la recherche, de former les différents acteurs aux NAM et à terme de mettre en œuvre des stratégies d'évaluation de risque de nouvelle génération afin de réduire l'expérimentation animale⁸⁷. Il vise à réunir les parties prenantes, autant publiques que privées (institutions académiques, agences de réglementation, évaluateurs de risques et industriels), autour de débats réguliers dans l'élaboration de stratégies d'amélioration et de feuilles de route pour le futur de l'évaluation des risques des substances chimiques pour la santé et l'environnement²⁰.

D'autres projets européens tels que l'EPAA (« European Partnership for Alternative approaches to Animal Testing ») travaillent au regroupement des savoirs entre scientifiques de l'industrie et la Commission Européenne pour collaborer sur des études de cas impliquant les NAM⁸⁸. Évaluant de façon critique les difficultés rencontrées, ce projet permet d'identifier les priorités pour l'inclusion des NAM dans la réglementation : une participation plus importante des autorités réglementaires, un élargissement des possibilités et des secteurs d'utilisation des NAM, et une augmentation des besoins en termes de formation sur les différentes méthodes⁸⁸.

À travers différentes incitations réglementaires, les gouvernements s'accordent sur la suppression progressive des tests sur les animaux, matérialisant leur volonté sur diverses feuilles de route :

- Aux États-Unis, l'Environmental Protection Agency (EPA) a déclaré en 2020 vouloir réduire de 30 % les études sur les animaux d'ici à 2025, et les supprimer d'ici 2035⁸⁹ ;
- Au Canada, la loi Bill S-5 (2022) vient remplacer l'ancienne loi pour la protection de l'environnement (Canadian Environmental Protection Act (CEPA), 1999), luttant également pour la suppression totale des tests animaux d'ici à 2035⁹⁰ ;
- En Europe, l'Union européenne promeut la suppression des tests animaux, avec en 2003 l'interdiction de l'utilisation d'animaux pour les tests cosmétiques (EU Directive 76/768/EEC⁹¹), jusqu'au dernier groupe de travail organisé en décembre 2023 par la Commission Européenne, réunissant différentes autorités réglementaires pour traiter de la réduction des tests animaux⁹².

Les méthodes NAM ayant le potentiel d'accélérer la production et l'interprétation de données toxicologiques, il est aujourd'hui nécessaire de les inclure dans les réglementations. Les différents acteurs et textes de lois impliqués évoluent, mais des difficultés persistent quant à l'implémentation de ces nouvelles méthodes.

3. Défis et perspectives pour une meilleure intégration

Il est aujourd'hui nécessaire d'avoir des méthodes validées et globalement harmonisées afin de réduire significativement les tests sur les animaux. En effet, des méthodes validées au niveau mondial renforceront la confiance et permettront dans les faits une acceptation réglementaire.

Des outils communs entre les diverses réglementations mondiales permettront de gagner en clarté, augmenter la comparabilité des données et de gagner du temps en construisant des familles de substances pouvant être évaluées ensemble et construire des bases de données de référence favorisant le développement des méthodes alternatives⁸⁴. Cependant, la réglementation exige un niveau de preuve équivalent à celui des données animales et constitue donc un frein à l'acceptation réglementaire des méthodes alternatives. Les méthodes non animales manquent aussi de certification, à cause d'un nombre insuffisant de substances évaluées *in silico*, la reproductibilité qui ne peut être garantie, des modèles d'interprétation des données qui ne sont pas établis, des essais qui ne sont pas disponibles à l'échelle commerciale et du manque de données de référence⁸⁴. Or, il est aujourd'hui nécessaire de définir les exigences réglementaires nécessaires à la certification des NAM. En effet, le nombre important de nouveaux essais disponibles conduit à une validation uniquement basée sur les performances en termes de prédiction, quand les lignes directrices

pourraient s'appuyer davantage sur les critères définissant cette performance et les contextes d'utilisation réelle des tests²⁰.

En termes d'avancées en toxicologie prédictive, certaines méthodes alternatives sont maintenant reconnues par la réglementation en Europe. Le règlement REACH encadre désormais les méthodes alternatives les plus utilisées : le read-across, la méthode QSAR et les méthodes *in vitro* pour la toxicité cutanée et oculaire (irritation, corrosion, sensibilisation)⁹³. Ces méthodes seront développées dans les parties suivantes.

Selon l'OCDE, les travaux à prioriser pour avancer sur les méthodes alternatives concernent le développement et la validation de nouvelles méthodes non animales pour les effets toxiques sur le système nerveux, le système immunitaire ainsi que les perturbateurs endocriniens, mais aussi des approches prédictives de l'exposition interne⁸⁴. De plus, le rapport de l'ECHA publié en Novembre 2023 définit les problématiques scientifiques actuelles posant les défis réglementaires les plus urgents⁹⁴. Parmi ces défis se trouvent la protection contre les produits chimiques les plus dangereux, la pollution chimique de l'environnement, l'amélioration de la disponibilité des données chimiques, mais aussi l'arrêt de l'expérimentation animale.

La recherche sur les outils NAM ne cessant d'avancer, les perspectives futures en vue d'améliorer ces méthodes sont nombreuses :

- Plutôt que de tester des hypothèses, il serait intéressant de développer des méthodes alternatives capables de générer elles-mêmes ces hypothèses.
- L'optimisation de la validation et de la reproductibilité des résultats par des contrôles expérimentaux efficaces permettraient d'améliorer la confiance dans les observations.

- Les avancées en cours quant à la puissance informatique devraient conduire à des résultats plus pertinents et des simulations représentatives des systèmes biologiques et physicochimiques.
- La prise en compte d'un volume de données *in vivo* plus important, provenant de nombreux échantillons, de taille plus importante, permettrait d'identifier de nouvelles pistes de recherche.
- La généralisation des résultats serait permise par la prise en compte d'un échantillon plus large lors des prédictions informatiques, et permettrait de se rapprocher des conditions réelles d'application, conduisant à des prédictions plus réalistes⁷⁹.

Combiner les résultats obtenus grâce aux méthodes NAM avec ceux provenant des expérimentations traditionnelles s'impose aujourd'hui comme un aspect essentiel aux progrès de la recherche en toxicologie et à la prise de décision réglementaire en évaluation des risques.

II. Évaluation de la pertinence des données NAM dans l'approche d'évaluation du danger : études de cas sur la toxicologie orale aiguë

A. Introduction

1. Contexte

Certains paramètres toxicologiques spécifiques comme la corrosion et l'irritation de la peau, les lésions oculaires ou la mutagénicité sont souvent évalués à l'aide d'approches non animales. Mais pour d'autres paramètres comme la toxicité systémique aiguë (voie orale, cutanée ou inhalation), la toxicité par administration répétée, la toxicité pour la reproduction, la cancérogénicité et la neurotoxicité, l'évaluation de la sécurité des produits chimiques repose encore largement sur l'expérimentation animale⁹⁵. La présente étude se concentrera sur l'utilisation d'alternatives à l'expérimentation animale dans la toxicité systémique aiguë par voie orale, la modélisation *in silico* l'ayant déjà réalisé auparavant⁹⁶.

2. Définition des méthodes utilisées

a. Read-across

Pour obtenir des données sur les substances chimiques, le « read-across » est une approche alternative qui peut aider à réduire les essais sur les animaux, ce qui permet d'économiser du temps et des coûts pour le développement des produits⁹⁷. Le read-across est défini comme « une méthode alternative permettant de prédire la toxicité d'un composé par analogie structurelle avec d'autres composés dont la toxicité est mieux connue »⁴ : une prédiction d'un critère toxicologique pour une substance cible, avec peu ou pas de données, est

faites sur la base de l'identification de substances analogues et qui bénéficient d'un nombre de données conséquent et adéquat⁹⁸. Simple et transparente, cette méthode nécessitant un jugement d'experts est appréciée dans la prise de décision institutionnelle, dans le cadre de l'évaluation de la sécurité. En effet, elle ne nécessite pas de connaissances approfondies en statistiques et modélisation moléculaire, parfois nécessaires pour interpréter d'autres modèles de prédiction⁹⁹.

Toutefois, pour appréhender l'utilisation du read-across et son acceptabilité au niveau réglementaire, il est nécessaire d'en définir et d'évaluer les incertitudes existantes. En effet, certains biais peuvent avoir un impact sur l'acceptation réglementaire des prédictions effectuées en read-across, notamment à cause d'une sélection incorrecte de substance(s) ou d'étude(s) source(s)¹⁰⁰. D'autres incertitudes peuvent être liées à la qualité et à la pertinence des données collectées sur les effets toxicologiques, la connaissance des mécanismes conduisant à la toxicité (à la fois chimique et biologique), l'argumentaire du read-across et la justification de la similarité. En effet, un groupe chimique fonctionnel pouvant se retrouver dans de nombreux composés, cela peut conduire à une grande sensibilité dans la prédiction, mais à une faible spécificité^{101,102}.

Afin de surmonter ces biais, et lorsque les données fournies par le read-across sont insuffisantes, des approches complémentaires reposant sur les relations quantitatives entre structure chimique et activité biologique (QSAR) sont mises en place pour la prédiction, afin de soutenir l'évaluation de la toxicité.

b. QSAR

La modélisation QSAR est une approche informatique reposant sur des techniques statistiques et d'apprentissage automatique afin d'établir une corrélation entre les propriétés chimiques d'une molécule et les effets toxicologiques mesurés¹⁰³. Elle prend en compte un ou plusieurs descripteurs de

la structure moléculaire, un effet à prédire et la relation mathématique qui permet lier la structure moléculaire avec l'effet toxique¹⁰⁴. La méthode QSAR est largement utilisée pour fournir des informations fiables sur la toxicité des produits chimiques, une validation rigoureuse du modèle étant nécessaire pour obtenir des modèles QSAR robustes et prédictifs¹⁰⁵.

De vastes ensembles de données expérimentales, soumises à l'évaluation statistique¹⁰⁶, permettent de fournir des prédictions quantitatives (QSAR) ou qualitatives (SAR) sur la toxicité des composés. Ils sont devenus un outil majeur dans l'évaluation de la sécurité chimique et dans l'aide à la décision réglementaire depuis l'adoption en 2006 de la législation « Enregistrement, évaluation, autorisation et restriction des produits chimiques » (REACH) par l'Union européenne¹⁰⁷. Les outils Web pour la modélisation QSAR comprennent des bases de données robustes publiques, telles que CATMoS (« Collaborative Acute Toxicity Modeling Suite »), OPERA (« Open (Quantitative) Structure-activity/property Relationship App »)¹⁰⁸, MultiCASE (« Multiple Computer Automated Structure Evaluation »)¹⁰⁹, la base de données QSAR du JCR¹¹⁰ et la base de données (Q)SAR danoise (Danish (Q)SAR Database¹¹¹). En 2008, QSAR Toolbox, un outil rassemblant différents modèles QSAR validés, a été développé par l'OCDE pour faciliter et appliquer des approches par catégories, définir des hypothèses mécanistes et dériver et appliquer des modèles QSAR adaptés basés sur ces hypothèses¹¹².

Il existe de nombreux modèles (Q)SAR permettant de prédire un large éventail de paramètres qualitatifs et quantitatifs, mais tous n'ont pas été validés par rapport aux données expérimentales dans des contextes spécifiques. Pour ces contextes spécifiques, il reste à vérifier si les résultats reflètent réellement l'expérimentation *in vivo* : ainsi, les données QSAR ne peuvent être utilisées à des fins d'identification des dangers et d'évaluation des risques que dans le cadre

d'une approche fondée sur le poids de la preuve (après une évaluation précise des informations dérivées du modèle)¹¹³.

3. Toxicité orale aiguë

La toxicité orale aiguë décrit les effets systémiques indésirables d'une substance à la suite d'une exposition orale unique sur une période de 24 heures. Elle se concentre sur la détermination de la dose létale ou « DL₅₀ », définie comme la dose unique de substance censée provoquer la mort de 50 % des animaux d'un groupe expérimental¹¹⁴.

Les essais de toxicité orale aiguë sont requis par la réglementation REACH pour les substances ayant des effets potentiellement mortels à la suite d'une exposition aiguë par voie orale, et afin de caractériser la toxicité et d'attribuer des catégories de danger. L'évaluation de la toxicité orale aiguë permet de fournir, sur la base d'une relation dose-réponse, des informations relatives aux restrictions ou précautions appropriées à prendre lors de la manipulation des substances, du transport ou de l'utilisation, et ainsi permettre l'étiquetage de la substance¹¹⁵. Les critères de classification sont basés sur des « valeurs seuils » spécifiques, déterminées grâce à des études suivant des méthodes d'essai certifiées et incorporées dans le règlement de l'UE¹¹³ ou selon les lignes directrices de l'OCDE.

Comme indiqué dans le règlement européen n° 1272/2008 (CLP), les catégories de danger pour la toxicité aiguë sont basées sur des données animales, tandis que la toxicité spécifique d'organes cibles après une exposition unique est basée sur des preuves provenant d'humains ou bien d'animaux de laboratoire¹¹⁶. Les exigences de la réglementation REACH, décrites dans le guide de l'ECHA, s'appliquent en toxicité systémique aiguë pour les substances fabriquées ou importées dans l'Union européenne à raison d'une tonne ou plus par an (spécifiées dans les annexes VII et VIII)¹¹⁷. Pour classer les substances présentant

une toxicité orale aiguë, le CLP de l'Union européenne a adopté une approche à quatre niveaux, de la catégorie 1 à la catégorie 4, mais en considérant également une catégorie 5, non toxique. Les différentes catégories sont définies comme suit :

- Catégorie 1 : $DL_{50} \leq 5$ mg/kg de poids corporel (pc) (mortel en cas d'ingestion)
- Catégorie 2 : $5 < DL_{50} \leq 50$ mg/kg pc (mortel en cas d'ingestion)
- Catégorie 3 : $50 < DL_{50} \leq 300$ mg/kg pc (toxique en cas d'ingestion)
- Catégorie 4 : $300 < DL_{50} \leq 2000$ mg/kg pc (nocif en cas d'ingestion)
- Catégorie 5 : $DL_{50} > 2000$ mg/kg pc (non toxique)¹¹⁸

Les catégories CLP sont établies en comparant les prédictions expérimentales de DL_{50} aux intervalles de chaque catégorie. La DL_{50} est exprimée dans les réglementations en mg/kg de poids corporel (mg/kg pc), mais pour les études *in silico*, elle est généralement exprimée en concentration molaire (c'est-à-dire en mol/kg p.c.), car c'est le nombre de molécules présentes dans le produit plutôt que leur poids qui est responsable des effets toxiques observés¹¹⁹.

Les études *in vivo* de toxicité aiguë se basent sur la concentration létale et sont donc fortement critiquées, tant sur le plan scientifique que sur le plan éthique^{120,121}. En effet, les tests de toxicité aiguë restent les plus pratiqués aujourd'hui¹²² et résultent donc en une grande utilisation d'animaux. En réponse à ces problématiques, les lignes directrices de l'OCDE concernant les essais ont progressivement évolué pour s'aligner sur les principes des 3R, qui consistent à minimiser le risque de douleur et de détresse pour les animaux utilisés dans la recherche¹²³.

Certaines approches permettant d'éviter de recourir à des espèces animales présentent un intérêt croissant dans le cadre de la plupart des évaluations réglementaires de toxicité aiguë, selon les instances réglementaires de l'UE. Cependant, malgré de nombreux efforts, le passage des tests de toxicité orale

aiguë à des approches *in silico* est contraint par une compréhension insuffisante des nombreuses voies de toxicité conduisant à une toxicité systémique aiguë¹²⁴.

4. Objectifs de recherche

L'objectif de ce travail est, pour différentes substances analogues, de comparer les valeurs expérimentales *in vivo* aux données de DL₅₀ prédites à l'aide de deux approches NAM : le read-across et la méthode QSAR. L'objectif est de montrer la pertinence de ces différentes approches en matière de prédiction de DL₅₀ pour la toxicité orale aiguë et ainsi déterminer si les approches NAM permettent de modifier la classification CLP et donc, par conséquent, l'évaluation des risques. Ceci devrait permettre d'objectiver les perspectives en termes d'acceptabilité réglementaire d'une stratégie utilisant des données *in silico*.

Grâce au logiciel QSAR Toolbox, des données ont été extraites de la « Base de données sur la toxicité orale aiguë » (« Acute Oral Toxicity Database »¹²⁵) et trois groupes de produits chimiques structurellement similaires ont été créés. Ensuite, des données expérimentales provenant de tests de toxicité orale aiguë *in vivo* ont été rassemblées pour ces substances, à partir d'études respectant les lignes directrices de l'OCDE. Deux types d'approches par read-across, utilisant des substances analogues présentes dans chaque groupe, ont été réalisées sur le logiciel. La première approche consistait en la prédiction de la DL₅₀ basée sur la moyenne des valeurs de DL₅₀ des plus proches voisins, et la seconde a permis de calculer des prédictions basées sur la valeur minimale de DL₅₀ des voisins les plus proches. Les données QSAR ont été recueillies dans la base de données en ligne « Danish (Q)SAR Database »¹¹¹. Les principes de classification CLP ont été utilisés pour classer les substances dans différentes catégories réglementaires, sur la base des résultats fournis par les trois approches de prédiction. Les valeurs expérimentales et les résultats de prédiction ont été comparés en termes de

classification dans les catégories CLP, afin de conclure sur la pertinence des approches NAM dans l'appui à la décision réglementaire (figure 5).

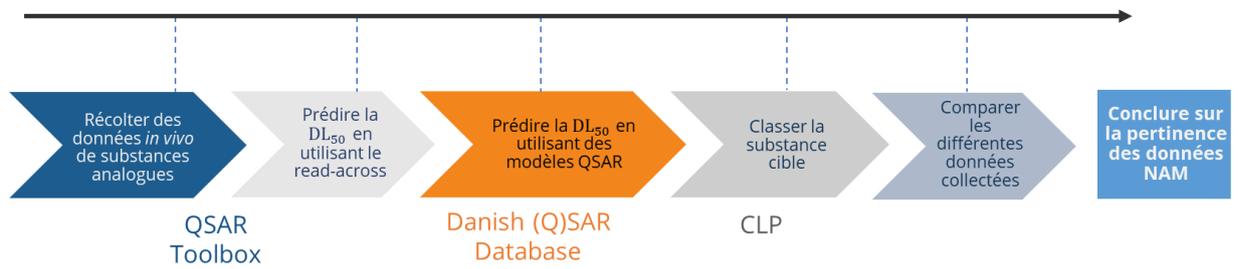


Figure 5: Schéma de la prédiction et de l'évaluation de la toxicité orale aiguë à l'aide d'outils *in silico*.

B. Matériel et méthodes

1. Outils utilisés

a. QSAR Toolbox

Ce travail vise à prédire la toxicité orale aiguë de substances chimiques *in silico*, en utilisant une plateforme innovante : QSAR Toolbox¹¹². QSAR Toolbox est un logiciel développé par l'Organisation de coopération et de développement économiques (OCDE), qui œuvre pour la promotion des politiques publiques améliorant le bien-être économique et social partout dans le monde¹²⁶. QSAR Toolbox a été conçu pour réduire l'utilisation d'animaux dans les tests de laboratoire, le coût de ces tests, et augmenter le nombre de produits chimiques évalués pour leurs effets sur la santé humaine et l'environnement¹²⁷.

Les principales caractéristiques comprennent l'identification des caractéristiques structurelles et du mode d'action d'un produit chimique cible, l'utilisation de catégories prédéfinies pour identifier les produits chimiques présentant les mêmes caractéristiques structurelles ou le même mécanisme et l'utilisation des données expérimentales existantes. Cette approche vise à combler le(s) manque(s) de données à l'aide d'approches telles que le read-across, une analyse des tendances (« trend analysis ») ou des modèles QSAR¹²⁷.

Il existe six fonctionnalités majeures dans QSAR Toolbox, correspondant aux six modules individuels disponibles dans le système, qui peuvent être utilisés séquentiellement ou indépendamment, afin de rassembler et d'évaluer les données expérimentales et prédites disponibles¹²⁸ (figures 6 et 7).

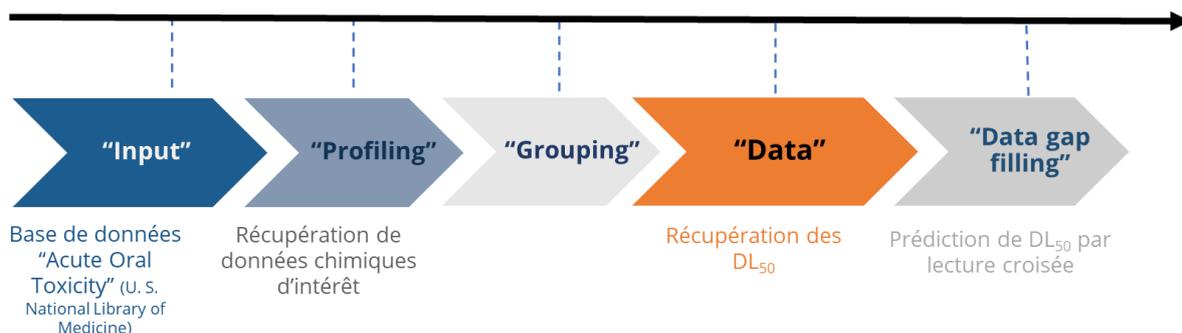


Figure 6 : Étapes conduisant à la prédiction d'une donnée, sous QSAR Toolbox.

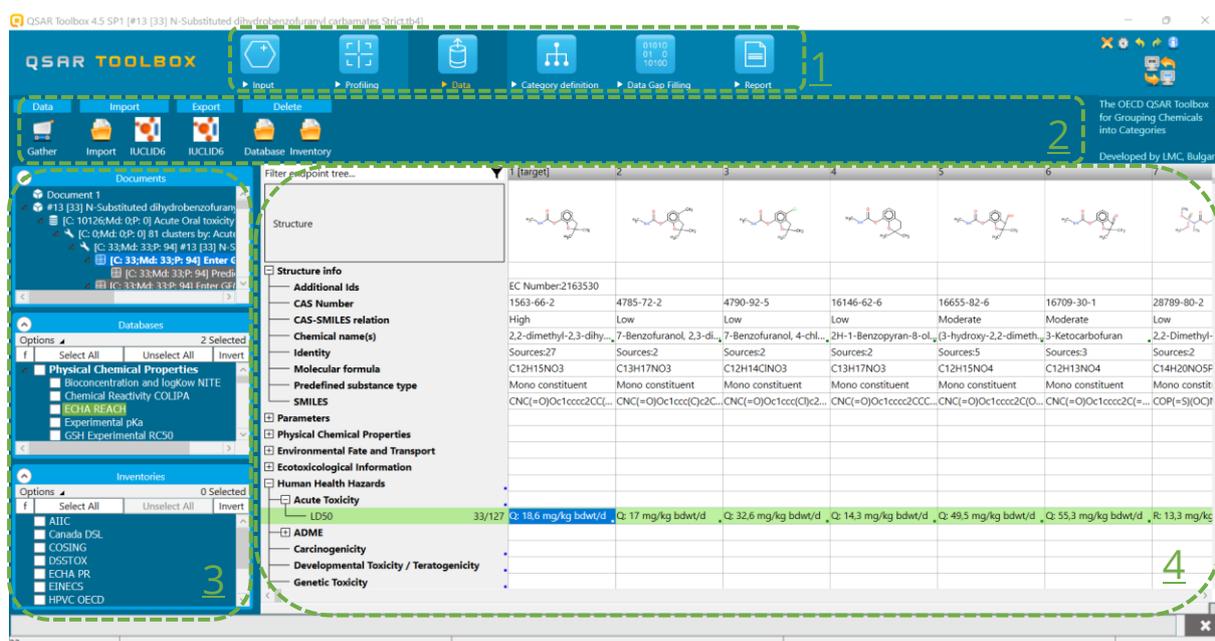


Figure 7 : Les différents composants de l'interface de QSAR Toolbox : les différents modules (1), les principales fonctionnalités de chaque étape (2), les options à sélectionner (3) et la matrice de données (4).

Pour les besoins de ce travail, seuls les quatre modules suivants utilisés seront présentés : « input », « profiling », « data » et « data gap filling ».

- « Input »

Ce module offre à l'utilisateur plusieurs moyens de saisir le produit chimique d'intérêt, par exemple par le numéro unique d'enregistrement unique auprès de la banque de données du Chemical Abstracts Service (CAS), le nom (trivial, commercial, IUPAC, etc.), la sélection à partir d'une liste prédéfinie de produits chimiques ou d'une base de données.

Étant donné que toutes les fonctions ultérieures sont basées sur la structure chimique, l'objectif est de s'assurer que la structure moléculaire attribuée au produit chimique cible est la bonne. La bonne attribution de l'identité chimique est évaluée qualitativement comme « élevée », « modérée » ou « faible ».

- « Profiling »

Ce module fait référence au processus de récupération des informations pertinentes (caractéristiques structurales et mécanistiques) sur le composé cible, stockées dans la base de données, permettant d'attribuer cette substance à des catégories prédéfinies contenant des substances chimiques similaires (figure 8) et ainsi soutenir la prédiction basée sur les catégories¹²⁹.

- Définition du critère d'évaluation à étudier et collecte des données (onglet « Data »)

Le module « Data » fait référence au processus électronique de récupération des données de toxicité et sur le devenir dans l'environnement qui sont stockées dans la base de données. Les données sont réparties en quatre sections : propriétés physico-chimiques, devenir et transport dans l'environnement, informations écotoxicologiques et risques pour la santé humaine.

- Comblent les lacunes en matière de données (« Data gap filling »)

En fonction de la situation, il est nécessaire de choisir le mécanisme le plus pertinent pour combler les lacunes dans les données. Le logiciel propose trois approches pour combler ce manque de données : la read-across, l'analyse des tendances (« trend analysis ») et une bibliothèque de modèles QSAR indépendants, qui peuvent être appliqués à la substance cible. La lecture croisée et l'analyse des tendances peuvent être visualisées graphiquement, la variable d'intérêt en ordonnée, et une variable indépendante en abscisse, cette dernière étant une caractéristique appropriée s'appliquant aux membres de la catégorie.

b. Danish (Q)SAR database

La « Danish (Q)SAR database » est une base de données accessible au public de plus de 650 000 substances chimiques et qui contient des informations sur les effets, notamment les propriétés physico-chimiques, le devenir dans l'environnement, la bioaccumulation, l'écotoxicité, l'absorption, le métabolisme et la toxicité¹³⁰. Elle permet de réaliser des prédictions basées sur des modèles de QSAR. Cette base de données a été développée en 2004 par des autorités européennes et danoises, sous l'égide de l'Université Technique du Danemark (« Technical University of Denmark » (DTU)), le « DTU National Food Institute », avec le soutien de l'Agence danoise de protection de l'environnement (« Environmental Protection Agency »), du Conseil nordique des ministres (« Nordic Council of Ministers ») et de l'Agence européenne des produits chimiques (ECHA). Son objectif principal est de soutenir la prise de décision réglementaire et l'évaluation des risques liés aux produits chimiques. Cette plateforme contient des informations sur un large éventail de produits chimiques, y compris les pesticides, les produits pharmaceutiques et les produits chimiques industriels¹¹¹.

Danish (Q)SAR database est accessible en ligne et les résultats sont mis à disposition gratuitement *via* un système de recherche avancée comprenant une recherche par nom chimique, numéro CAS (US Chemical Abstract Service), structure chimique ou similarité.

L'utilisateur peut choisir entre une recherche par identifiant de la molécule, soit par son nom, soit en le cherchant dans une liste, provenant d'une base de données REACH ou autre. Cet identifiant peut être sous la forme d'un numéro CAS, d'un numéro CE (numéro de la Communauté européenne) soit par l'identificateur de composés PubChem. La recherche de la substance peut également être effectuée grâce à la structure chimique ou le système SMILES (Simplified Molecular Input

Line Entry System). Lorsque la substance correcte apparaît, il est possible de récupérer un document Word avec les données QSAR.

2. Approches

a. Collecte de données

La base de données sur la toxicité orale aiguë (« Acute Oral Toxicity Database (AOT DB) ») a été jugée pertinente pour rassembler les substances associées à une toxicité orale aiguë *in vivo*. Cette base de données publique ne contient que des données évaluées par des experts, dont les informations expérimentales sont extraites de ChemIDplus, appartenant à la plateforme de la National Library of Medicine des États-Unis¹³¹. La base de données comprend 10 128 produits chimiques, avec des données de DL₅₀, obtenues après exposition à des rats pendant 24 heures. Les données fournies sont extraites de plus de 10 000 références bibliographiques provenant d'une sélection d'articles de recherche. La recherche a été effectuée pour le critère « DL₅₀ », en rassemblant les données expérimentales de DL₅₀ pour toutes les substances, dans la « Base de données sur la toxicité orale aiguë ».

b. Création de groupes de substances

Afin de créer un groupe primaire d'analogues proches en termes de structure chimique et d'activité, la base de données « Acute Oral Toxicity Database » a été séparée en groupes de substances. Ces substances, extraites de cette base de données, ont en effet été regroupées grâce au système de profilage de QSAR Toolbox en fonction de leurs similitudes en termes de structure chimique, de modes d'action et de métabolisation.

Ce mode de sélection permet un regroupement structurel pour identifier les analogues ayant la même toxicité orale aiguë, c'est-à-dire la même réactivité ainsi que la même réponse toxicologique^{96,132}.

Ainsi, 31 groupes de similarité structurelle ont été mis en évidence, et trois ont été choisis pour être analysés : le groupe des « carbamates » (carbamates dihydrobenzofuranyl N-substitués), le groupe des « phosphorothioates » (diethyl alkyl aryl phosphorothioates) et le groupe des « organophosphorés » (alkyl et/ou aryl ester organophosphorés à longue chaîne).

c. Prédiction

Read-across

Afin d'obtenir une prédiction *in silico* de DL₅₀ pour chacune des substances des groupes sélectionnés, la méthode du read-across a d'abord été utilisée.

Différentes options de calcul existent pour approximer la prédiction par read-across : la médiane, la médiane inférieure, la médiane supérieure, le minimum, le maximum, la moyenne et la moyenne géométrique. Dans ce travail, deux options de calcul ont été utilisées et comparées. Chaque substance est devenue à son tour une substance cible, les autres analogues étant utilisés comme substances sources, ce qui a permis de faire des prédictions pour l'ensemble du groupe.

La première approche de calcul était basée sur la moyenne des valeurs DL₅₀ obtenues expérimentalement des analogues les plus proches, ainsi appelée « read-across moyen ». Le but a donc été de faire la moyenne de toutes les valeurs de DL₅₀ obtenues *in vivo* disponibles, provenant de différentes sources, pour chaque substance analogue, afin d'obtenir une DL₅₀ par substance analogue. Les substances analogues les plus proches en termes structurels et mécanistiques seront ensuite utilisées pour conclure sur une prédiction par read-across pour la substance cible.

La deuxième approche de calcul a utilisé les plus faibles valeurs de DL₅₀ obtenues expérimentalement des substances analogues. La valeur prédite pour la substance cible provient ainsi de la valeur minimale de DL₅₀ des substances analogues les plus proches. Elle a ainsi été appelée « read-across minimal ». La

première approche est celle choisie par défaut dans QSAR Toolbox, tandis que l'autre approche est considérée comme un scénario « catastrophe », une approche plus protectrice car les valeurs analogues les plus toxiques sont choisies pour la prédiction.

Le read-across est visualisé graphiquement avec en ordonnée l'effet considéré (ici $\log(DL_{50})$) en fonction d'une caractéristique (descripteur) des membres de la catégorie (ici $\log(Kow)$). Comme il existe une association statistique avec les données toxicologiques à prédire, le choix du descripteur est souvent automatiquement fait par le logiciel¹³³. En effet, dans le cas de la toxicité orale aiguë, il a été prouvé qu'une approche d'évaluation des risques peut être pertinente en utilisant le coefficient de partage octanol/eau (Kow) comme descripteur¹³⁴.

QSAR

Les données QSAR sont extraites à l'aide de Danish (Q)SAR Database, en renseignant le numéro CAS de chaque substance. Une fois que la substance d'intérêt apparaît dans la section « résultats », un rapport QSAR peut être téléchargé, fournissant un fichier au format .RTF contenant toutes les prédictions, basées sur trois modèles QSAR indépendants (CASE Ultra, Leadscope Predictive Data Miner et SciQSAR¹³⁵). Parmi les prédictions concernant les propriétés physico-chimiques, l'(éco)toxicité et le devenir dans l'environnement, on retrouve celles liées à la toxicité humaine, dont la toxicité orale aiguë. Ces dernières proviennent de tests sur des rats auxquels le composé a été administré par voie orale. Elles ont été compilées par ADC/Labs à partir du Registre des effets toxiques des substances chimiques (« Registry of Toxic Effects of Chemical Substances » ou RTECS) et de l'interface informatique IUCLID (« International Uniform Chemical Information Database »), servant à saisir, stocker et gérer les données intrinsèques aux substances chimiques¹³⁶. La compilation réalisée par ADC/Labs

permet donc d'obtenir un résultat QSAR numérique, accessible dans le rapport fourni par Danish (Q)SAR Database.

Ce modèle QSAR fournit également un indice de fiabilité (RI), un nombre entre 0 et 1, pour donner une évaluation quantitative de la fiabilité. La base de données danoise (Q)SAR ne fournit des prédictions que dans le domaine d'applicabilité du modèle (défini par un RI supérieur à 0,5). Ce dernier détermine la fiabilité du modèle : il est considéré comme la zone de l'espace chimique couverte par les substances et dans laquelle les prédictions sont valables¹³⁷.

d. Classification réglementaire en toxicité orale aiguë

L'attribution d'une catégorie CLP à une substance s'effectue en comparant la valeur prédite de la DL₅₀ aux valeurs seuils CLP, telles que décrites précédemment.

La performance de prédiction, c'est-à-dire la proportion de substances ayant la même catégorie CLP observée et prédite, a ensuite été évaluée. Ainsi, une étiquette a été attribuée à chaque substance de chaque groupe :

- « Classe équivalente » si la valeur prédite de la DL₅₀ classe la substance dans la même catégorie que celle attribuée expérimentalement ;
- « Surestimation » si la catégorie prédite attribuée est inférieure à la catégorie expérimentale, indiquant une surestimation du danger ;
- « Sous-estimation », si la catégorie prédite attribuée est supérieure à celle attribuée expérimentalement, indiquant une sous-estimation du danger.
- « Non classé », si la substance est prédite dans la catégorie 5 (non toxique) alors qu'elle est classée expérimentalement dans la catégorie 4.

Une fois les substances classées dans une catégorie en fonction des prédictions, la justesse est calculée pour évaluer la capacité de classification en utilisant la formule suivante :

$$Accuracy = \frac{Total_{category\ 1-5}}{Total_{predictions}} \times 100$$

où $Total_{category\ 1-5}$ est le nombre total de substances correctement prédites.

Les pourcentages de résultats correctement estimés, surestimés et sous-estimés, ainsi que les composés non classés sont représentés pour chaque groupe sous forme d'histogrammes. Les prédictions correctes correspondent donc à la justesse de la classification, les résultats surestimés et sous-estimés à la mauvaise classification des produits chimiques.

C. Résultats

Trois études de cas, sur trois groupes différents, ont été réalisées au cours de ce travail. Elles permettent d'illustrer le cadre utilisé pour l'évaluation de la fiabilité des données de read-across et des prédictions QSAR. Les données supplémentaires en annexe (tableaux S1, S2 et S3) correspondent aux détails sur les valeurs observées et prédites pour l'ensemble des données.

1. Étude de cas 1 : Carbamates

Les premières substances chimiques à évaluer font partie du groupe des carbamates (figure 8). Ces derniers appartiennent à la classe des insecticides et présentent un mécanisme d'action commun avec les organophosphorés : ils induisent la phosphorylation de l'acétylcholinestérase, mais de façon moins prolongée. Cela se traduit sur le plan toxicologique par une accumulation de neurotransmetteurs, en l'occurrence d'acétylcholine, au niveau du système nerveux et de la jonction neuromusculaire¹³⁸. Les différentes voies d'expositions (dermique, par inhalation ou par ingestion) entraînent des symptômes liés au système nerveux central (confusion, hallucinations, crises convulsives, etc.) et au système nerveux autonome (myosis, miction, bronchospasme, salivation, etc.). La gravité des symptômes varie en fonction de la classification de l'insecticide et de la dose à laquelle un être vivant est exposé¹³⁹.

Les substances à évaluer sont constituées de 33 molécules, considérés comme analogues et appartenant à la même catégorie (figure 9). En effet, en plus d'être rattachés au large groupe des carbamates, ils possèdent la caractéristique commune d'être substitués avec un ou plusieurs groupements dihydrobenzofuranyl.

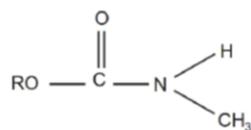


Figure 8 : Structure générale des carbamates, avec R = radicaux alkyles (adapté de Abd-Ella, 2011)¹⁴⁰

Filter endpoint tree...	1	2	3	...	32	33
Structure				...		
Structure info						
Additional ids	EC Number:2163530					
CAS Number	1563-66-2	4785-72-2	4790-92-5		1563-67-3	90293-54-2
CAS-SMILES relation	High	Low	Low		Low	Low
Chemical name(s)	2,2-dimethyl-2,3-dihy...	7-Benzofuranol, 2,3-di...	7-Benzofuranol, 4-chl...		Carbamic acid, methyl...	6,11-Dioxa-3-thia-2,4,...
Identity	Sources:27	Sources:2	Sources:2		Sources:1	Sources:1
Molecular formula	C12H15NO3	C13H17NO3	C12H14ClNO3		C11H13NO3	C21H29N5O7S
Predefined substance type	Mono constituent	Mono constituent	Mono constituent		Mono constituent	Mono constituent
SMILES	CNC(=O)Oc1cccc2CC(...	CNC(=O)Oc1ccc(C)c2C...	CNC(=O)Oc1ccc(Cl)c2...		CNC(=O)Oc1cccc2CC(...	CON=C(C)C(=NOC(=...

Figure 9 : Capture d'écran de QSAR Toolbox montrant le début et la fin du jeu de données avec les informations d'identification chimique et la structure des membres du groupe défini (groupe des carbamates dihydrobenzofuranyl N-substitués).

a. Résultats des profileurs

Afin d'identifier les caractéristiques chimiques spécifiques au critère étudié, plusieurs profileurs pertinents pour la toxicité orale aiguë ont été appliqués au groupe de substances (figure 10).

Filter endpoint tree...	1	2	3	4
Structure				
<input checked="" type="checkbox"/> Toxicokinetics, Metabolism and Distribution				
<input type="checkbox"/> Profiling				
<input type="checkbox"/> General Mechanistic				
iSafeRat® Mechanisms of toxic Action profiler	MechoA 6.1 Full MechoA: Full MechoA: >> Mec... 6: direct docking disru... 6: direct docking disru...	MechoA 6.1 Full MechoA: Full MechoA: >> Mec... 6: direct docking disru... 6: direct docking disru...	MechoA 6.1 Full MechoA: Full MechoA: >> Mec... 6: direct docking disru... 6: direct docking disru...	MechoA 6.1 Full MechoA: Full MechoA: >> Mec... 6: direct docking disru... 6: direct docking disru...
Protein binding by OASIS	Acylation Acylation >> Direct ac... Acylation >> Direct ac...			
Toxic hazard classification by Cramer	High (Class III)	High (Class III)	High (Class III)	High (Class III)
<input type="checkbox"/> Endpoint Specific				
Acute Oral Toxicity	N-Substituted dihydro...	N-Substituted dihydro...	N-Substituted dihydro...	N-Substituted dihydro...
<input type="checkbox"/> Empiric				
Organic functional groups	Aryl Benzofuran/Dihydrob... Carbamate Coumaran	Alkyl (hetero)arenes Alkyl-, alkenyl- and alk... Aryl Benzofuran/Dihydrob... Carbamate Coumaran	Aryl Aryl halide Benzofuran/Dihydrob... Carbamate Coumaran	Aryl Carbamate Dihydrochromene/ben...

Figure 10 : Capture d'écran de QSAR Toolbox montrant les résultats du profilage des 4 premières substances du groupe.

La première étape du profilage a été de comparer et d'évaluer les mécanismes d'action des substances chimiques présentes dans le cluster. Pour cela le profileur iSafeRat Mechanisms of toxic Action (ou MechoA) v1.1.2 (inclus dans QSAR Toolbox en mai 2023), qui permet de classer les produits chimiques en fonction de leur mécanisme d'action toxicologique chez différentes espèces, a été utilisé. Il constitue le point de départ de l'AOP, décrivant les événements moléculaires initiateurs (MIE), ainsi que, dans certains cas, les événements moléculaires clés qui suivent. Il a été développé à partir de plus de 1800 molécules, dont les données sont extraites de 350 publications, se basant sur des études mécanistiques expérimentales (principalement sur les mammifères et les poissons) et l'examen des structures chimiques.

Ainsi, le schéma de classification MechoA utilise un arbre de décision pour diviser les produits chimiques en 6 sous-groupes généraux de mécanisme d'action toxique¹⁴¹. Dans ce cas, le sous-groupe général auquel appartient le cluster est le

MechoA 6 « Direct docking interaction : binding to a docking site of a key protein (enzyme, receptor, ion channel) » (interaction directe avec un site d'amarrage moléculaire (une protéine clé telle qu'une enzyme, un récepteur ou un canal ionique) et la description complète correspond au MechoA 6.1 : « Acetylcholinesterase inhibition for animals and probably other esterases inhibition for all species » (inhibition de l'acétylcholinestérase pour les animaux et probablement d'autres estérases, pour toutes les espèces). Cela correspond au mécanisme général des carbamates, tel que décrit dans la littérature.

La deuxième étape du profilage a consisté à vérifier que les analogues de la catégorie se lient bien à une protéine, en utilisant des systèmes de reconnaissance des caractéristiques structurales (ou « alertes structurales »). Le système « Protein binding by OASIS », un des profileurs utilisés dans QSAR Toolbox, identifie si une substance contient une des 101 alertes, au sein de 11 types de mécanismes différents, responsables de l'interaction avec les protéines¹²⁸. Ainsi, les analogues sont cohérents en ce qui concerne la « liaison aux protéines par OASIS » : ils présentent tous une alerte pour une acylation directe (c'est-à-dire l'introduction d'un groupe acyle dans un composé) impliquant un groupe partant. Les carbamates sont connus pour être thermiquement instables, ce qui conduit à une dégradation de la substance, entraînant une réaction d'acylation qui diminue la toxicité orale¹⁴².

La « classification des dangers toxiques selon Cramer » est l'un des profileurs mécanistiques qui traite de la similarité entre les substances en termes de toxicité. Il s'agit de l'approche la plus connue permettant d'évaluer le seuil de préoccupation toxicologique (en anglais « Threshold of Toxicological Concern », ou TTC), qui vise à établir une valeur numérique du seuil d'exposition humaine pour toutes les substances chimiques en deçà de laquelle les risques toxicologiques pour l'homme seraient considérés comme négligeables¹⁴³. Ce seuil de

préoccupation toxicologique permet de définir trois classes de toxicité selon Cramer :

- Classe I : TTC = 30 µg/kg pc/jour (substances faiblement toxiques),
- Classe II : TTC = 9 µg/kg pc/jour (substances avec une toxicité intermédiaire),
- Classe III : TTC = 1,5 µg/kg pc/jour (substances hautement toxiques)¹⁴⁴.

Ici, les substances du groupe étudié appartiennent toutes à la classe III selon la classification des dangers toxiques de Cramer.

En ce qui concerne la cohérence en termes de groupes fonctionnels organiques, responsables des principales propriétés chimiques, les substances analogues du cluster présentent certaines similitudes en termes de groupements chimiques. Il est ainsi possible de retrouver au sein des 33 substances chimiques des groupements aryles et carbamates, et toutes les substances, sauf une, partagent les fonctions benzofurane/dihydrobenzofurane et coumarine.

En résumé, le processus de regroupement décrit est cohérent avec les principes de read-across, où les analogues sont rassemblés sur la base de leurs caractéristiques similaires¹²⁹. Dans le cas présent, les 33 substances du groupe sont similaires en ce qui concerne les caractéristiques structurelles et mécanistiques de la toxicité aiguë par voie orale.

b. Performance de la prédiction CLP

Des prédictions par read-across ont été effectuées dans la section adéquate du logiciel, en utilisant les valeurs expérimentales disponibles (figure 11).

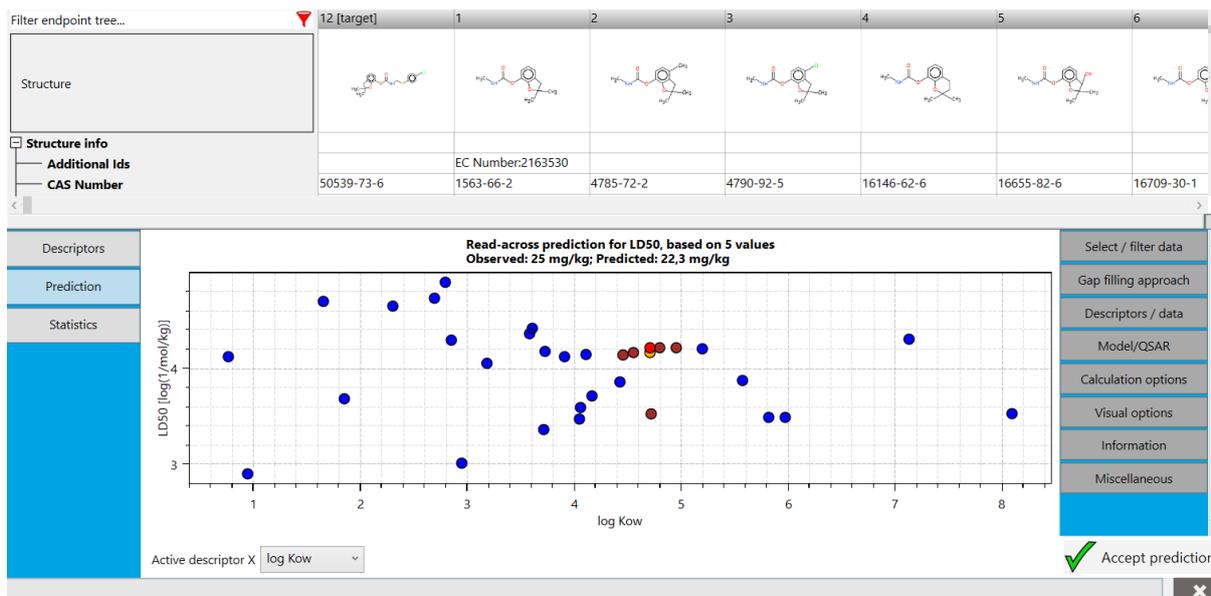


Figure 11 : Capture d'écran d'une prédiction par read-across, pour une substance chimique du groupe. Le graphique résultant montre les données expérimentales de tous les analogues par rapport à leur descripteur (log Kow). En rouge figure la valeur prédite de la DL₅₀ pour la substance cible, en marron les résultats expérimentaux disponibles pour les analogues du cluster utilisés et en bleu les résultats expérimentaux des analogues non utilisés.

L'étude de la performance de ces trois approches *in silico* (read-across moyen, read-across minimal et QSAR) est essentielle pour répondre à l'objectif de cette étude, à savoir la classification selon les critères de la réglementation CLP. Les résultats obtenus pour chacune des 33 substances sont regroupés dans le tableau S1.

Les résultats montrent que les trois approches *in silico* conduisent à une grande proportion de classification CLP correcte (« correct category » sur le graphique), respectivement 67 %, 43 % et 50 % en utilisant le QSAR, le read-across moyen (« read-across average ») et le read-across minimal (« read-across minimum ») (figure 12). En général, lorsque la classification est mauvaise, la prédiction a tendance à être surestimée (« overpredicted category ») plutôt que sous-estimée (« underpredicted category ») : 23 % de surestimation contre 10 % de sous-estimation, 37 % de surestimation contre 20 % de sous-estimation et 50 % de surestimation contre aucune sous-estimation, respectivement pour QSAR, le read-across moyen et le read-across minimal. Cela implique qu'en toxicité orale

aiguë, le danger a tendance à être surestimé, en particulier avec le read-across utilisant les valeurs minimales de DL₅₀.

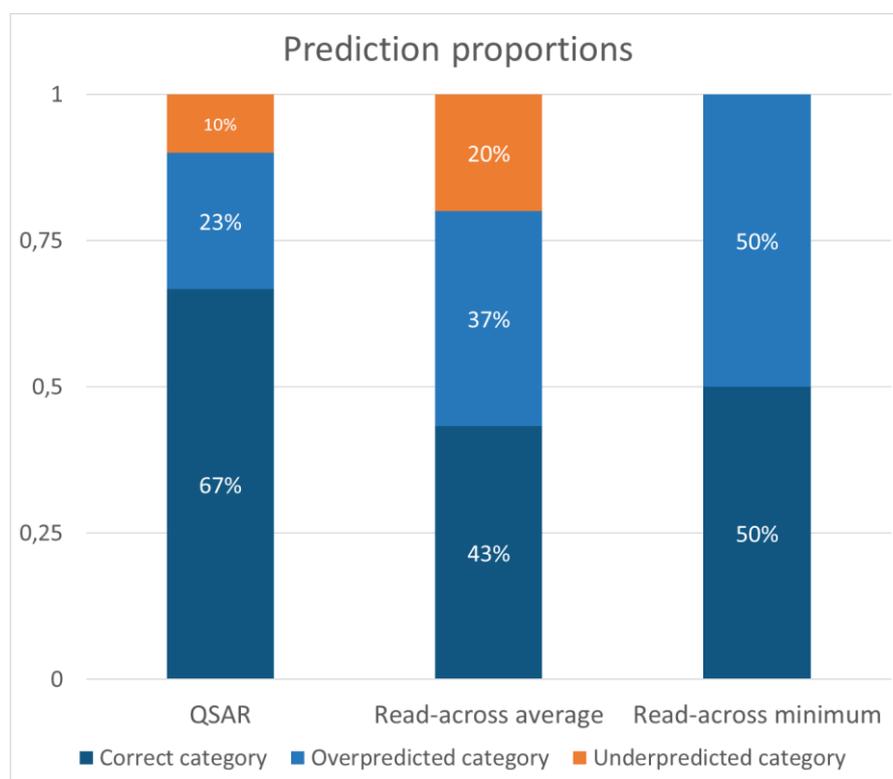


Figure 12 : Résumé des performances de prédiction de DL₅₀ en pourcentages, pour cette première étude de cas comprenant 33 substances chimiques. La classification des données *in silico* (approches QSAR et read-across) est effectuée selon les critères de catégories CLP, et comparée aux valeurs expérimentales. Les pourcentages représentent le nombre de catégories prédites par rapport aux catégories attribuées expérimentalement.

En termes de performance par catégorie CLP (figure 13), les résultats ont révélé des performances de prédiction variées selon les trois approches *in silico*. Pour les composés de la catégorie 1, QSAR et read-across moyen ont sous-estimé dans 100 % des cas la catégorie correspondante expérimentalement, alors que la méthode du read-across minimal a permis de classer correctement 100 % des substances. Pour les composés de la catégorie 2, des classifications correctes ont été obtenues dans 100 %, 81 % et 75 % des cas en utilisant respectivement le QSAR, la read-across moyen et le read-across minimal. Des sous-estimations dans les prédictions ont été obtenues à 19 % par le read-across moyen et à 25 % par le read-across minimal. Enfin, la méthode QSAR a permis 64 % de surestimations quant aux substances de la catégorie 3, tandis que les deux approches de read-

across ont conduit à 100% de surestimation pour les composés de la catégorie 3. Cette méthode a également conduit à 36 % de classification correcte, pour les composés de la catégorie 3.

Dans l'ensemble, les composés de catégorie 3 ont majoritairement fait l'objet d'une surestimation quant à la prédiction de leur catégorie CLP. Les meilleures prédictions ont été obtenues par read-across, en utilisant les valeurs minimales de DL₅₀, avec uniquement des catégories correctement estimées ou surestimées. Ainsi, les résultats suggèrent que les prédictions les plus précises sont obtenues en utilisant le read-across avec des valeurs minimales de DL₅₀.

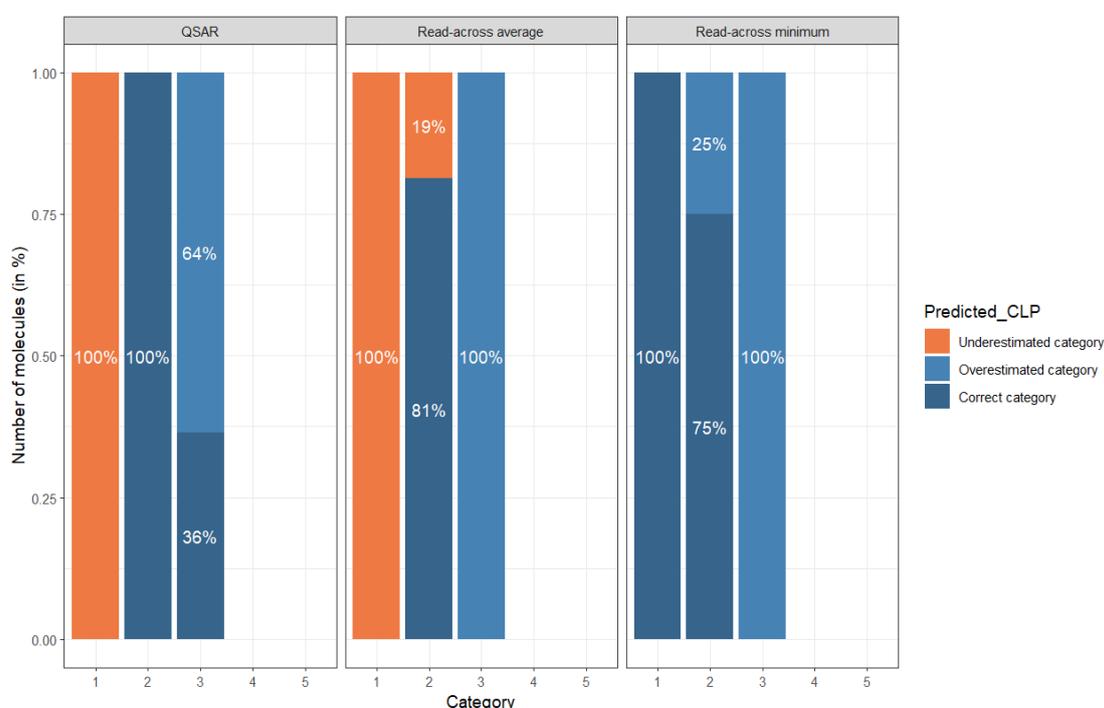


Figure 13 : Performance en termes de prédiction pour les trois approches par catégorie, comparée aux valeurs expérimentales. Les 33 valeurs de DL₅₀ obtenues *in vivo* dans ce groupe de substances couvrent les trois premières catégories. Le pourcentage des prédictions correctes, surestimées ou sous-estimées, obtenues *in silico*, pour chaque catégorie du cluster, sont affichées dans l'histogramme.

2. Étude de cas n° 2 : Phosphorothioates

Cette deuxième étude de cas se concentre sur un autre groupe, créé à la suite de la catégorisation par l'« Acute Oral Toxicity Database ». Il comprend des molécules appartenant au groupe des « phosphorothioates » (ou « esters

phosphorothioïques »). Ces composés, contenant des atomes de phosphore et de soufre, représentent une classe importante en raison de leur large éventail d'applications possibles, tant en agrochimie, comme les pesticides organophosphorés, qu'en chimie médicale¹⁴⁵. Ils se caractérisent par une apparition tardive des manifestations toxiques, en raison de leur activation en phosphorodithioates nécessaire aux effets cliniques¹⁴⁶.

Ce groupe de phosphorothioates, plus précisément « diéthyl alkyl aryl phosphorothioates », et dont la structure est présentée figure 15, contient 29 substances chimiques (figure 14) considérées comme analogues.

Filter endpoint tree...	1	2	3	...	28	29
Structure						
Structure info						
Additional Ids	EC Number:2208644	EC Number:2002717	EC Number:2025645		20195-07-7	39624-86-7
CAS Number	2921-88-2	56-38-2	97-17-6		Low	Low
CAS-SMILES relation	High	High	High		Phosphorothioic acid...	DOWCO 275
Chemical name(s)	Chlorpyrifos	ethyl parathion	Dichlofenthion		Sources:1	Sources:3
Identity	Sources:36	Sources:28	Sources:9		C8H19O3PS	C9H13FNO3PS
Molecular formula	C9H11Cl3NO3PS	C10H14NO5PS	C10H13Cl2O3PS		Mono constituent	Mono constituent
Predefined substance type	Mono constituent	Mono constituent	Mono constituent		CCOP(=S)(OCC)OC1nc(...	CCOP(=S)(OCC)OC(C)CCOP(=S)(OCC)Oc1cc...
SMILES	CCOP(=S)(OCC)OC1nc(...	CCOP(=S)(OCC)OC1cc...	CCOP(=S)(OCC)OC1cc...			

Figure 14 : Capture d'écran de QSAR Toolbox montrant le début et la fin de la matrice de données avec les informations d'identification chimique et la structure des membres du groupe défini (groupe « Diethyl aryl phosphorothioates »).

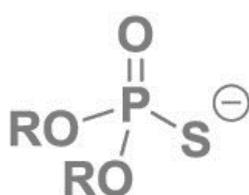


Figure 15 : Structure générale des phosphorothioates¹⁴⁶.

a. Résultats des profileurs

Pour ce groupe, le profileur MechoA correspondant est le numéro 6 « Direct docking interaction : binding to a docking site of a key protein (enzyme, receptor, ion channel) » (interaction directe avec un site d'amarrage moléculaire (une protéine clé telle qu'une enzyme, un récepteur ou un canal ionique)) et la

description complète correspond au MechoA 6.1 : « Acetylcholinesterase inhibition for animals and probably other esterases inhibition for all species » (inhibition de l'acétylcholinestérase pour les animaux et probablement d'autres estérases, pour toutes les espèces). En ce qui concerne la liaison avec les protéines comme définie par le profileur OASIS : tous les analogues présentent une alerte de substitution nucléophile (réaction SN2) sur les esters de phosphate ainsi qu'une acylation suivie d'une phosphorylation. Les substances faisant partie de ce groupe appartiennent toutes à la classe III selon la classification de Cramer, ainsi considérées comme des composés hautement toxiques ¹⁴⁴. En termes de groupements fonctionnels organiques caractéristiques, tous les composés chimiques partagent une fonction thiophosphate et tous, sauf cinq, possèdent une fonction aryle. D'autres fragments se retrouvent de façon sporadique chez d'autres analogues.

Ainsi, toutes les substances chimiques du groupe présentent des similitudes structurelles et mécanistiques au sein de la catégorie des phosphorothioates, conformément aux principes de read-across pour la toxicité orale aiguë.

b. Performance de la prédiction CLP

Le processus de prédiction de DL₅₀ utilisant le read-across a été effectué, suivi de l'application des principes de classification CLP aux valeurs prédites. Les résultats obtenus pour chacune des 29 substances sont regroupés dans le tableau S2.

Les résultats ont ainsi montré que 52 %, 48 % et 22 % des substances ont été prédites dans les bonnes catégories, en utilisant respectivement QSAR, le read-across moyen et le read-across minimal (figure 16). La surestimation a concerné 4 % des prédictions QSAR, 26 % des prédictions par read-across moyen et 67 % des prédictions par le read-across minimal. Enfin, 44 %, 26 % et 11 % des composés

ont été sous-estimés en utilisant respectivement les approches QSAR, read-across moyen et read-across minimal.

Parmi les trois approches, la méthode QSAR est responsable de la majorité des résultats sous-estimés, quand la méthode du read-across minimal a mené à la plus grande proportion de résultats surestimés. Bien que cette dernière approche ait produit moins de résultats corrects que les deux autres, elle a permis de réduire efficacement le nombre de catégories sous-estimées, ce qui en fait une méthode plus fiable et plus conservatrice pour minimiser le risque de sous-estimation du danger.

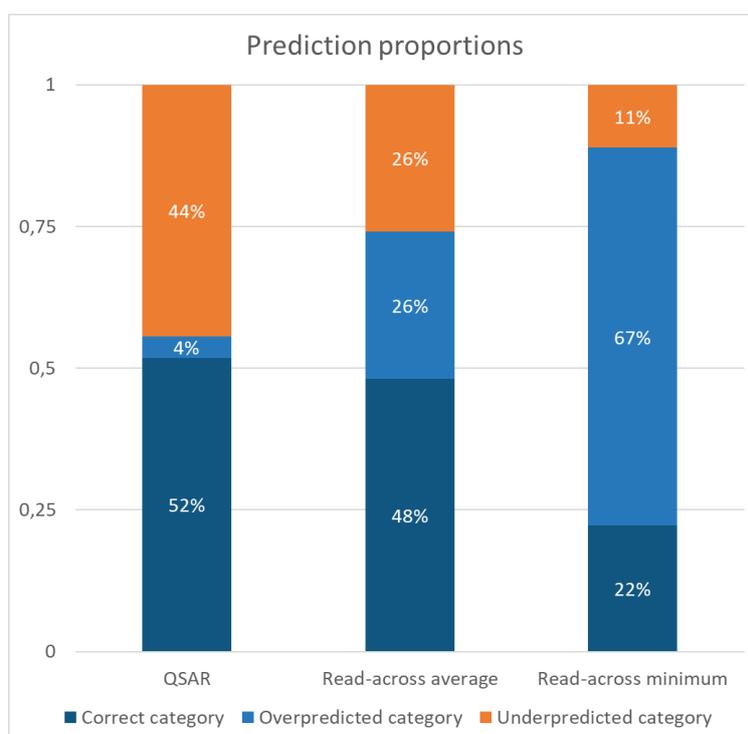


Figure 16 : Résumé de performances en termes de prédiction du DL50 en pourcentages, pour la deuxième étude de cas comprenant 29 substances chimiques. La classification des données *in silico* (approches QSAR et read-across) est effectuée selon les catégories CLP et comparée aux valeurs obtenues *in vivo*. Les pourcentages représentent le nombre de catégories prédites par rapport aux catégories attribuées expérimentalement.

Dans ce groupe, les substances appartiennent à des catégories allant de 1 à 4, selon les résultats expérimentaux (figure 17). Lors de l'utilisation d'approches *in silico*, la toxicité des composés de catégorie 1 ont tendance à être sous-estimés,

notamment 50 % d'entre eux en utilisant l'approche QSAR et 100 % des composés de la catégorie 1 en utilisant les deux approches de read-across. Pour ce qui est des substances de la catégorie 2 et 3, respectivement 67 % et 33 % ont été sous-estimées en utilisant l'approche QSAR. Les méthodes de read-across moyen et minimal ont conduit à une sous-estimation des résultats pour 42 % et 8 % des composés de la catégorie 2, respectivement. Pour ce qui est des catégories prédites à l'aide de QSAR, 50 %, 33 %, 67 % et 75 % pour les catégories 1, 2, 3 et 4, respectivement, sont correctes. En utilisant le read-across moyen, 58 % et 67 % des composés des catégories 2 et 3, ont fait l'objet d'une prédiction correcte, et 50 % des composés de la catégorie 2 ont également été prédits dans la bonne catégorie, en utilisant le read-across minimal.

Les résultats ont été sous-estimés pour 50 %, 67 % et 33 % des composés des catégories 1, 2 et 3, respectivement, en utilisant la méthode QSAR. Les deux approches de read-across ont conduit à une sous-estimation des résultats pour 100 % des composés de la catégorie 1. Pour les composés de la catégorie 2, 42 % ont été sous-estimés avec l'approche de read-across moyen et 8 % avec l'approche de read-across minimal.

Enfin, seuls les composés de la catégorie 4 ont vu leurs résultats surestimés en utilisant le QSAR (25 %). En utilisant le read-across moyen, 33 % des composés de la catégorie 3 ont été surestimés, alors que tous les composés de la catégorie 4 ont été surestimés (100 %). Le read-across minimal a conduit à une surestimation de 42 % des composés de la catégorie 2 et de tous les composés (100 %) des catégories 3 et 4.

Dans l'ensemble, les résultats des performances par catégorie ont montré que, en utilisant le read-across, les catégories CLP les plus nocives (c'est-à-dire les catégories 1 et 2) avaient tendance à être sous-estimées. En outre, plus la catégorie est élevée, plus les résultats ont tendance à être surestimés. Par rapport

aux valeurs expérimentales, la prédiction QSAR a conduit à la plus grande proportion de prédictions correctes, avec néanmoins plus de 50 % de résultats sous-estimés pour les catégories 1 et 2.

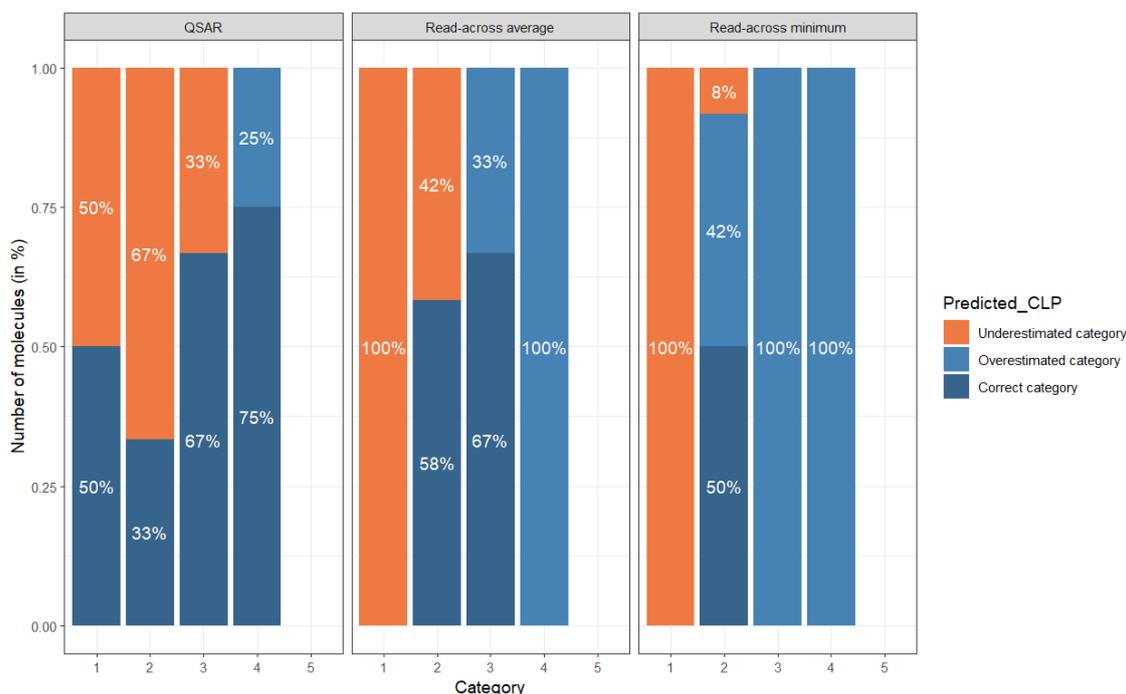


Figure 17 : Performance de prédiction selon les trois approches, par catégorie. Le groupe des phosphorothioates comprend 29 substances chimiques, couvrant quatre catégories. Les pourcentages de résultats corrects, surestimés et sous-estimés sont affichés dans ce graphique par catégorie de toxicité CLP, en utilisant les résultats des prédictions QSAR, read-across moyen et read-across minimal, par rapport aux catégories dérivées expérimentalement.

3. Étude de cas n° 3 : Esters organophosphorés

Les esters organophosphorés (OPE) (figure 18) sont dérivés de l'acide phosphorique, et sont principalement utilisés dans la fabrication de peintures, de teintures textiles, de produits électroniques et bien d'autres encore, afin d'augmenter la résistance aux flammes¹⁴⁷. A l'instar des deux groupes précédents, ces composés sont dérivés du groupe des organophosphorés, et partagent également la particularité d'être d'emblée actifs¹⁴⁸. Ils ont été conçus pour remplacer les polybromodiphényléthers, interdits en 2004 en raison de leur persistance dans l'environnement, de leur bioaccumulation et de leur toxicité¹⁴⁹.

Depuis, l'utilisation des retardateurs de flamme à base d'esters organophosphorés a augmenté et les préoccupations concernant l'exposition humaine se font croissantes. En effet, ces composés sont libérés dans l'environnement (par volatilisation ou abrasion), induisant leur présence dans l'air, l'eau potable et les aliments¹⁵⁰. Étant donné l'omniprésence des OPE, le risque pour la santé doit être examiné de manière plus approfondie par des études toxicologiques et épidémiologiques.

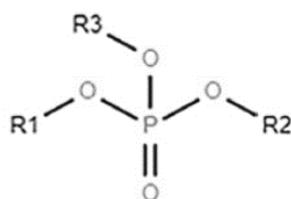


Figure 18 : Structure générale des esters organophosphorés.

Sélectionné dans la base de données « Acute Oral Toxicity Database », ce groupe d'OPE, avec pour particularité commune de posséder une chaîne longue d'alkyle, comprend 24 composés (figure 19).

Filter endpoint tree...	1	2	3	...	23	24
Structure						
Structure info						
Additional Ids	EC Number:2011035	EC Number:2011166	EC Number:2011229			
CAS Number	78-30-8	78-42-2	78-51-3		4281-67-8	6163-73-1
CAS-SMILES relation	High	High	High		Low	Low
Chemical name(s)	O-TOLYL PHOSPHATE	Phosphoric acid, tris(2-...	tributoxyethyl phosph...		Cyclohexyl phenyl pho...	Ethanol, 2-methoxy-...
Identity	Sources:17	Sources:31	Sources:27		Sources:1	Sources:1
Molecular formula	C21H21O4P	C24H51O4P	C18H39O7P		C18H21O4P	C9H21O7P
Predefined substance type	Mono constituant	Mono constituant	Mono constituant		Mono constituant	Mono constituant
SMILES	Cc1cccc1OP(=O)(Oc1...	CCCC(C)COP(=O)(O...	CCCCOCCOP(=O)(OCC...		O=P(OC1CCCCC1)(Oc...	COCCOP(=O)(CCOC)...

Figure 19 : Capture d'écran de QSAR Toolbox montrant le début et la fin de la matrice de données, avec les informations d'identification chimique et la structure des membres de ce groupe d'esters organophosphorés.

a. Résultats des profileurs

En appliquant le profileur de mécanismes d'action toxique iSafeRat, 16 composés apparaissent comme MechoA 6, plus précisément MechoA 6.1, tels que décrits précédemment. Le reste des substances appartient au MechoA 2 : « Enzymatic hydrolysis: a mixture of both direct accumulation and enzymatic

hydrolysis generating acidity » (pro-déstabilisation par hydrolyse enzymatique) , dont la description complète est la suivante : « MechoA 2.1 : Hydrolysis to destabilizers : direct membrane destabilization and enzymatic hydrolysis, giving only narcotic products » (hydrolyse vers des déstabilisateurs : déstabilisation directe des membranes et hydrolyse enzymatique, ne générant que des produits narcotiques¹⁵¹). Ce mécanisme d'action par hydrolyse enzymatique s'explique par le mécanisme interne des inhibiteurs d'acétylcholinestérases, il s'agit en effet de la principale voie de détoxification chez les mammifères pour éliminer ces composés toxiques dans l'urine ¹⁵².

Selon OASIS, tous les analogues présentent une alerte pour la substitution nucléophile (réaction SN2) sur les esters de phosphate ainsi qu'une acylation suivie d'une phosphorylation. Les substances faisant partie de ce groupe appartiennent toutes à la classe III selon la classification de Cramer, impliquant que ce sont tous des composés hautement toxiques¹⁴⁴. En termes de fragments structurels caractéristiques, tous les produits chimiques partagent la fonction ester de phosphate, avec différents autres groupes retrouvés en minorité dans les substances analogues.

Ainsi, en ce qui concerne les résultats des profileurs, toutes les substances chimiques du groupe présentent des similitudes structurelles et mécanistiques avec les esters organophosphorés, conformément aux principes de lecture croisée pour la toxicité orale aiguë.

b. Performance de la prédiction CLP

Les résultats ont montré que le read-across utilisant les valeurs moyennes de DL₅₀ a conduit à la plus grande proportion de classifications correctes (74 %), malgré 26 % des substances chimiques ayant été classées comme non toxiques (figure 20). La read-across minimal a donné lieu à la plus grande proportion de prédictions surestimées (70 %). Cette approche a permis de classer correctement

26 % des substances et de prédire que 4 % d'entre elles n'étaient pas toxiques. Enfin, l'approche QSAR a permis de classer 21 % des produits chimiques comme non toxiques, et le reste des substances ont été prédites soit dans la bonne catégorie (52 %), soit dans une catégorie inférieure (26 %).

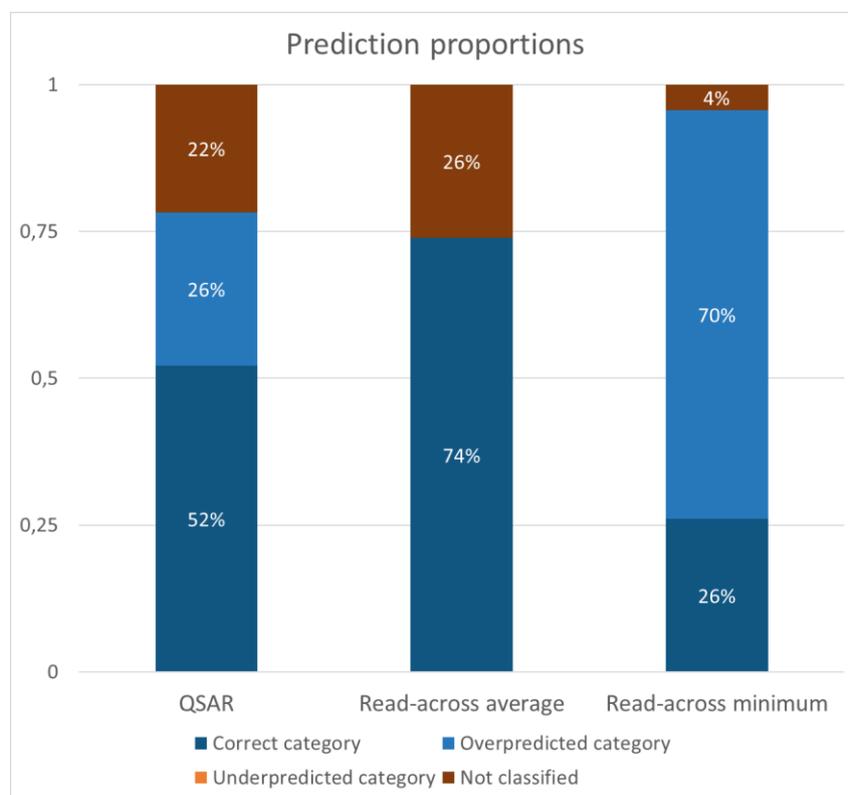


Figure 20 : Performance des prédictions selon les trois approches, avec les différentes proportions de catégories correctement estimées, les catégories surestimées et les catégories sous-estimées, par rapport au nombre total de substances chimiques du cluster. La proportion « not classified » correspond aux composés initialement classés comme toxiques mais qui ont été prédits en catégorie 5, c'est-à-dire une catégorie non toxique. Cela constitue ainsi une autre situation de sous-estimation de la catégorie.

En ce qui concerne la prédiction par catégorie, la méthode QSAR a classé correctement 17 % des composés de la catégorie 4 et 65 % des composés non toxiques (catégorie 5). Le read-across par la moyenne a classé correctement 100 % des substances non toxiques, et le read-across minimal, 83 % des composés de la catégorie 4 et 6 % des composés de la catégorie 5. La plus grande proportion de surestimation a été obtenue principalement lors de l'utilisation du read-across minimal : 94 % des composés non toxiques ayant été prédits comme toxiques. En outre, en utilisant QSAR, 35 % des composés non toxiques ont été

prédits comme toxiques. Le read-across utilisant la moyenne des valeurs analogues n'a quant à lui pas produit de surestimations.

Enfin, les résultats de prédiction par catégorie ont montré que le danger était sous-estimé pour la plupart des substances appartenant à la catégorie 4, en utilisant QSAR (83 %) et le read-across moyen (100 %). Toutefois, en utilisant la méthode du read-across minimal, 17 % des composés de la catégorie 4 ont été sous-estimés. Au sein du cluster, la plupart des substances chimiques se trouvent dans la quatrième catégorie, c'est-à-dire proches de la valeur limite de toxicité (2000 mg/kg p.c.), ce qui signifie qu'une sous-estimation des résultats les placera en catégorie 5, la catégorie non toxique. Ainsi, la sous-estimation du danger des composés en catégorie 4 conduira à des composés « non classés » (figure 21). Le détail des résultats obtenus pour chacune des 24 substances est disponible dans le tableau S3.

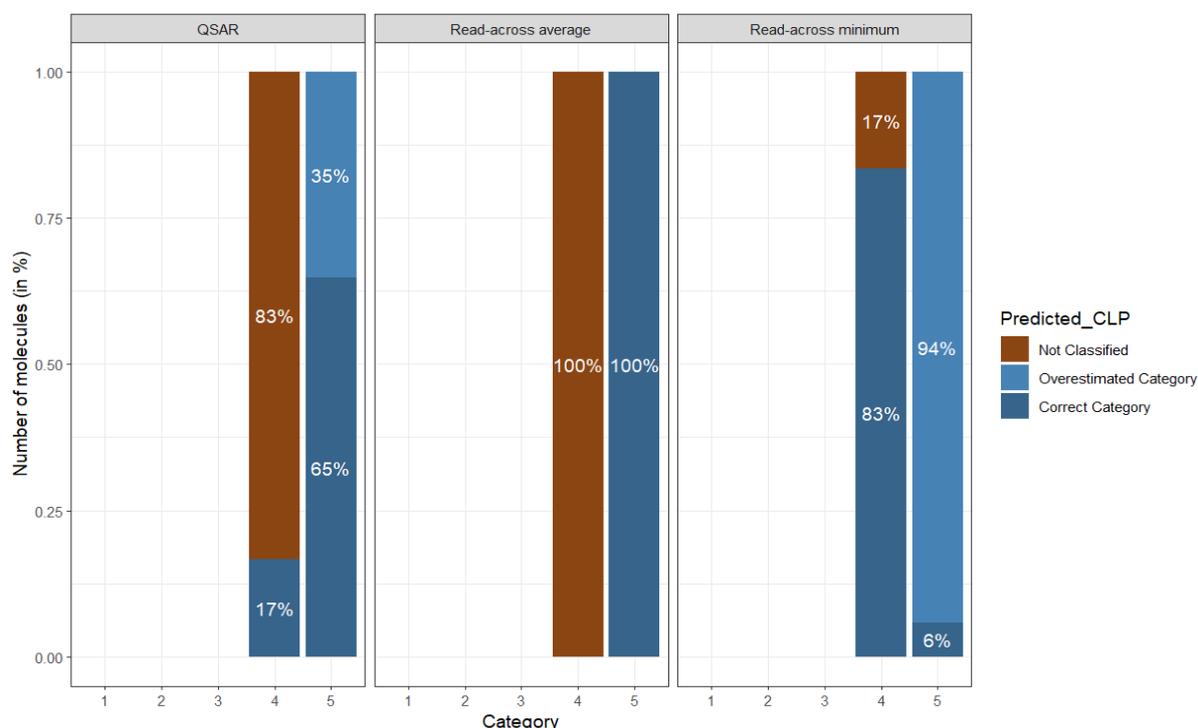


Figure 21 : Performance de prédiction selon les trois approches, par catégorie. Le groupe des esters organophosphorés comprend 24 substances chimiques, couvrant la catégorie CLP 4 ainsi que les substances non toxiques (catégorie 5). Les pourcentages de résultats corrects, surestimés et sous-estimés sont affichés dans ce graphique par catégorie de toxicité, en utilisant les résultats des prédictions QSAR, read-across moyen et read-across minimal, par rapport aux catégories dérivées expérimentalement.

4. Total des prédictions exactes

Le tableau suivant résume les prédictions correctes, pour chaque étude de cas, en utilisant chacune des trois approches (tableau 1). Il apparaît ainsi que l'approche *in silico* la plus précise pour prédire la toxicité orale aiguë est le QSAR, avec 57 % de prédictions correctes. Le read-across moyen induit plus de la moitié de résultats corrects, avec 55 % de prédictions correctes. L'approche de read-across par la valeur minimale est la moins juste des trois méthodes, avec 32,7 % de catégories correctement prédites.

Table 1: Total des prédictions. Les colonne « Total » donne une moyenne des résultats corrects, pour chaque méthode QSAR, read-across moyen et read-across minimal. Les résultats sont donnés en pourcentages (%) de résultats correctement prédits.

	Étude de cas 1	Étude de cas 2	Étude de cas 3	Total
QSAR	67 %	52 %	52 %	57 %
Read-across moyen	43 %	48 %	74 %	55 %
Read-across minimal	50 %	22 %	26 %	32.7 %

D. Discussion

Les méthodes *in silico* sont de plus en plus utilisées dans les industries chimiques afin de contribuer à l'évaluation de la toxicité des substances, depuis l'enregistrement jusqu'à l'étiquetage, en passant par la découverte précoce de médicaments. Bien que, d'un point de vue réglementaire, les études *in vivo* restent majoritaires dans l'évaluation de risques pour la toxicité orale aiguë, les méthodes alternatives sont de plus en plus impliquées dans les scénarios « weight-of-evidence ». En outre, l'identification de nouveaux mécanismes toxicologiques grâce aux systèmes *in silico* permet d'élargir leur utilisation dans un contexte réglementaire. Cependant, en raison du manque de standardisation de ces tests, il est difficile d'obtenir une acceptation plus large. Par conséquent, il est essentiel de renforcer la confiance dans leur capacité à prédire la toxicité humaine aiguë¹⁵³.

Cette étude réalisée sur trois groupes prédéfinis, partageant des caractéristiques structurelles et mécanistiques en toxicité orale aiguë, vise à donner un aperçu de la justesse de prédiction de trois approches *in silico* différentes, d'abord pour estimer les valeurs de DL₅₀ et, ensuite, pour classer ces substances dans les catégories de danger CLP appropriées liées au potentiel toxique aigu par voie orale. Les catégories de danger CLP attribuées selon les valeurs expérimentales et celles attribuées par des valeurs prédites ont été comparées, et la justesse calculée selon un pourcentage de substances prédites dans la bonne catégorie.

Ainsi, la justesse allait de 32,7 % en utilisant le read-across minimal à 57 % en utilisant QSAR. Les différences dans la conception des modèles pourraient expliquer les disparités en termes de résultats. En effet, les données QSAR sont basées sur un large ensemble de substances et sur la relation spécifique entre leurs variables structurelles (c'est-à-dire les paramètres structurels 2D ou 3D) avec un effet toxicologique défini, tandis que la méthode read-across est basée sur un

groupe plus restreint d'analogues similaires, choisis sur la base d'hypothèses de similarités toxicologiques structurelles.

Les groupes choisis étaient représentatifs des différentes catégories de danger CLP, couvrant un large éventail de valeurs de DL₅₀. Bien que les groupes ne soient pas similaires en termes de structure et de mécanismes, ils sont comparables en termes de résultats de prédiction. En effet, l'obtention d'une DL₅₀ plus élevée que lors des mesures expérimentales, c'est-à-dire une sous-estimation de la toxicité d'une substance, semble se produire davantage pour les catégories les plus toxiques de chaque groupe. Cela est probablement dû au fait que les fourchettes de toxicité aiguë dans les catégories de danger CLP sont plus larges dans les catégories les moins toxiques que dans les catégories les plus toxiques, induisant un biais lors de la classification (tableau 2).

Table 2 : Catégories CLP de toxicité. Les valeurs limites de DL₅₀ sont exprimées en mg/kg de poids corporel.

Catégorie	1	2	3	4	5
Valeurs limites CLP	DL ₅₀ ≤ 5	5 < DL ₅₀ ≤ 50	50 < DL ₅₀ ≤ 300	300 < DL ₅₀ ≤ 2000	DL ₅₀ > 2000

Il apparaît également que le read-across minimal, utilisant donc les valeurs minimales de DL₅₀ des substances analogues, est plus protecteur en termes de prédiction du danger. En effet, la plupart des valeurs prédites grâce à cette approche sont soit dans la bonne catégorie, soit surestimées par rapport aux valeurs expérimentales, dans tous les groupes. Cette approche, qui élimine le biais potentiel lié au jugement d'expert, semble minimiser le risque de classer des composés hautement toxiques dans une classe supérieure, c'est-à-dire les considérer comme plus sûrs. Ainsi, cette approche d'aide à la décision semble la plus protectrice parmi les deux read-across étudiés.

Le logiciel QSAR Toolbox fournit des outils nécessaires à la prédiction mais ne génère pas d'automatisation du flux de travail, ce dernier restant au choix de

l'utilisateur. Ainsi, outre les données provenant de Danish QSAR Database, les deux approches read-across fournissent une justification mécanistique du point de vue de l'utilisateur. Cela permet de conclure que ces prédictions *in silico* ne sauraient en l'état être utilisées comme éléments essentiels au processus décisionnel, mais plutôt comme une fonctionnalité d'aide à la décision.

Les incertitudes générées par le flux de travail utilisé pour la prédiction impliquent que le danger engendré par une substance peut être surestimé ou sous-estimé. Précédemment, il a été documenté que la classification d'une substance dans une catégorie adjacente est cohérente quant aux incertitudes qui accompagnent les valeurs de toxicité orale aiguë obtenues *in vivo*¹⁰⁸. En effet, il a été rapporté que les études *in vivo*, bien que suivant les lignes directrices des tests, contiennent une marge d'erreur inhérente aux modèles animaux ou à la méthode de test elle-même¹⁵⁴. Ces incertitudes limitent la capacité de prédiction des dangers, rendant une classification sans ambiguïté impossible pour de nombreuses substances. Chaque approche ayant ses propres limites, la synthèse des données *in vivo* et *in silico* semble être la meilleure solution. Cela semble être un moyen rationnel d'évaluer les dangers des substances, les données *in silico* servant principalement d'informations complémentaires par rapport aux données expérimentales. Les méthodes alternatives seraient utilisées en priorité, quand les tests *in vivo* seraient réservés aux substances prédites comme plus nocives.

En outre, le signalement des dangers sous forme d'étiquetage et de mention de danger selon la toxicité aiguë varie selon les classes de toxicité : les substances de classe 1, 2 et 3 sont étiquetées avec le pictogramme « danger », tandis que la classe 4 reçoit le pictogramme « avertissement » (tableau 3). Les mentions de danger pour les classes 1 et 2 sont « H300 : mortel en cas d'ingestion », tandis que la classe 3 est « H301 : toxique en cas d'ingestion » et la

catégorie 4, « H302 : nocif en cas d'ingestion ». Ainsi, cet étiquetage remet en question l'ambiguïté dans la classification entre les valeurs observées et prédites : la catégorisation d'une substance de la classe 1 à la classe 2 ne signifie pas une sous-estimation du danger puisque l'étiquetage ne change pas. Toutefois, une sous-estimation du danger se produira si une substance de classe 3 est prédite comme appartenant à la 4, étant donné que le pictogramme, le mot indicateur et la mention de danger sont différents, un cas qui ne s'est pas produit dans ce travail.

Table 3: Signalement des dangers par étiquetage pour la toxicité orale aiguë. Adapté de « Guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures », Version 5.0, 2017¹⁵⁵.

Éléments d'étiquetage relatifs à la toxicité orale aiguë				
Classification	1	2	3	4
Pictogramme				
Signal	Danger	Danger	Danger	Avertissement
Mention de danger	H300 : Mortel en cas d'ingestion	H300 : Mortel en cas d'ingestion	H301 : Toxique en cas d'ingestion	H302 : Nocif en cas d'ingestion

Ce travail, illustré par trois études de cas, visait donc à explorer les logiciels existants pour les outils *in silico* et comparer les résultats avec les données *in vivo*. Ainsi, après avoir constitué trois groupes de substances avec des similarités mécanistiques en ce qui concerne la toxicité orale aiguë, les approches « read-across » ont été utilisées pour effectuer des prédictions.

Le modèle utilisant le read-across minimal, a fourni une meilleure protection en termes de risque prévisionnel pour la toxicité aiguë que les autres, car il a souvent surestimé le danger. Le modèle QSAR a fourni les prévisions les plus justes, par rapport aux valeurs expérimentales. Il est donc important d'évaluer et de sélectionner l'approche la plus appropriée, ou une combinaison d'approches, en fonction du groupe de substances étudié. En outre, les approches *in silico*, et leurs

capacités de prédiction, sont censées répondre différemment à chaque critère d'évaluation toxicologique (c'est-à-dire mutagénicité, toxicité pour la reproduction et le développement, cancérogénicité, ...). Ici, la toxicité orale aiguë est le critère d'évaluation étudié, les conclusions soulignées dans cette étude s'appliquent donc spécifiquement à ce critère, et ne peuvent être valables pour d'autres effets sans une évaluation similaire de la qualité prédictive et de la pertinence des approches *in silico* utilisées.

Afin d'encadrer l'évaluation de la toxicité orale aiguë *in silico*, des outils de travail ont été développés tels que l'« Automated read-across workflow »⁹⁶, qui cherche à normaliser l'utilisation de QSAR Toolbox. Quoi qu'il en soit, les rapports de read-across générés automatiquement, étayés par des prédictions QSAR, doivent être davantage justifiés et validés par l'ECHA pour répondre aux exigences du règlement REACH. Enfin, de nouveaux modèles d'automatisation de read-across voient le jour, comme les modèles RASAR (« Read-Across Structure Activity Relationship »), combinant le « machine learning » avec les modèles de read-across, sur de larges bases de données¹⁸. Le read-across traditionnel est en effet basé sur la comparaison de quelques analogues chimiques, et un raisonnement au cas par cas pour chaque substance, mais est également dépendant d'une opinion d'expert. Ici, les modèles automatisés RASAR se basent sur l'apprentissage supervisé (« supervised learning » en anglais), qui utilisent des ensembles de données pour créer des algorithmes de prédiction, fournissant des modèles statistiques à partir d'informations sur les similarités chimiques. Ces nouveaux modèles de prédiction ont ainsi pu prouver leur efficacité, croisant de nombreuses données chimiques pour prédire les résultats de neuf des tests les plus couramment pratiqués sur les animaux à des fins de sécurité chimique. Les résultats ont montré 87 % de justesse pour 190 000 substances chimiques dont la classification était connue et 81 % de reproductibilité des tests suivant les lignes directrices de l'OCDE, pour chaque critère d'évaluation¹⁸.

D'autres travaux sont nécessaires pour améliorer la justesse de prédiction et accroître la confiance dans les résultats des NAM, en analysant davantage de groupes de substances, en identifiant des variables supplémentaires susceptibles d'influencer les résultats et en élargissant l'ensemble des critères d'évaluation pertinents à étudier. Il conviendrait également de définir une démarche permettant d'intégrer la réglementation CLP à l'utilisation de données *in silico*. Un atelier récent organisé par l'ECHA a mis l'accent sur la nécessité, pour les parties prenantes au sein de l'UE, d'accroître la coopération entre législations et l'implication dans les projets scientifiques impliquant les NAM. Par exemple, le projet collaboratif « Animal-free Safety assessment of chemicals : Project cluster for Implementation of novel Strategies » (ASPIS), récemment mis en œuvre dans le cadre de l'appel à projets NAM de la Commission européenne, représente la plus grande organisation de recherche financée par le secteur public en vue d'améliorer l'évaluation de la sécurité des substances chimiques sans recourir à l'expérimentation animale¹⁵⁶.

À la variété des outils NAM existants ou en cours d'élaboration s'ajoute la disponibilité des données, notamment grâce au criblage haut-débit d'un nombre important de molécules : par exemple le projet Tox21¹⁵⁷ a permis d'analyser 8000 substances au travers de plus de 85 tests de criblage. A cela s'ajoutent les diverses données médicales disponibles, les bases de données sur les structures chimiques ou encore les données présentes dans la littérature. Le nombre et la variété des données toxicologiques représentent un défi en termes de gestion de ces données, mais créent des opportunités pour l'intégration de méthodes alternatives utilisant l'intelligence artificielle. Sa capacité d'analyse de données à grande échelle en fait un outil indispensable à la toxicologie prédictive moderne.

E. Résumé des caractéristiques des substances

1. Étude de cas 1

Tableau S1 : Les 33 substances chimiques membres du groupe des carbamates sont référencées avec leur numéro CAS et leur nom de substance. Les valeurs expérimentales retrouvées, ainsi que les prédictions en read-across et QSAR sont indiquées pour chaque substance.

N°	CAS	Nom de la substance	Expérimental valeur (mg/kg pc)	Classification CLP des données expérimentale s	Valeur moyenne de lecture croisée (mg/kg pc)	Lecture croisée valeur minimale (mg/kg pc)	QSAR (mg/k g pc)
1	90293-54-2	Acide 8-((diméthylamino)carbonyl)-5-oxo-2,4,9-triméthyl-, 2,3-dihydro-2,2-diméthyl-7-benzofuranyl ester de 6,11-dioxa-3-thia-2,4,7,10-tétrazaadodéca-7,9-diénoïque	10,00	2	47,80	9,27	20,70
2	1563-66-2	Carbofuran	5,00	1	16,70	2,82	18,60
3	4785-72-2	7-Benzofuranol, 2,3-dihydro-2,2,4-triméthyl-, méthylcarbamate (9CI)	12,00	2	12,70	3,00	17,00
4	4790-92-5	7-Benzofuranol, 4-chloro-2,3-dihydro-2,2-diméthyl-, méthylcarbamate (9CI)	248,00	3	8,77	3,26	32,60
5	16146-62-6	Acide carbamique, méthyl-, ester de 2,2-diméthyl-8-chromanyle	3,00	1	16,80	4,40	14,30
6	16655-82-6	3-Hydrocarbofurane	18,00	2	17,60	4,44	49,50
7	16709-30-1	3-Ketocarbofuran	295,00	3	9,93	4,40	55,30
8	28789-80-2	2,2-Diméthyl-2,3-dihydrobenzofuranyl-7-N-diméthoxyphosphinothioyl-N-méthylcarbamate	150,00	3	26,90	13,30	22,34
9	37430-50-5	Acide carbamique, méthyl((trichlorométhyl)thio)-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	125,00	3	41,70	24,80	49,20
10	50539-66-7	Acide carbamique, méthyl(phénylthio)-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	58,70	22,00	25,30
11	50539-67-8	Acide carbamique, méthyl((4-méthylphényl)thio)-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	44,80	23,60	32,90
12	50539-70-3	Acide carbamique, ((p-tert-butylphényl)thio)méthyl-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	125,00	3	38,90	23,60	46,10
13	50539-73-6	Acide carbamique, ((4-chlorophényl)thio)méthyl-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	32,40	22,30	42,29
14	50539-74-7	Acide carbamique, ((p-bromophényl)thio)méthyl-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	35,50	25,00	42,17
15	50539-76-9	Acide carbamique, méthyl((pentachlorophényl)thio)-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	80,00	31,50	45,32
16	50539-80-5	Acide carbamique, méthyl(méthylthio)-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	5,00	1	17,70	3,41	23,60
17	50539-81-6	Acide carbamique, (éthylthio)méthyl-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	18,00	3,59	25,10
18	50673-06-8	Acide carbamique, méthyl((1,1,2,2-tétrachloroéthyl)thio)-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	125,00	3	27,80	25,70	94,68
19	50673-08-0	Acide carbamique, (butylthio)méthyl-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	60,10	3	45,00	22,30	30,30
20	50673-09-1	Acide carbamique, (dodécylthio)méthyl-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	125,00	3	50,00	21,00	110,00

21	50673-10-4	Acide carbamique, ((3,4-dichlorophényl)thio)méthyl-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	39,30	24,40	26,80
22	50673-11-5	Acide carbamique, ((2,5-dichlorophényl)thio)méthyl-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	39,30	24,40	26,80
23	50802-69-2	Acide carbamique, ((4-fluorophényl)thio)méthyl-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	61,70	26,40	23,10
24	50802-70-5	Acide carbamique, méthyl((4-nitrophényl)thio)-, ester de 2,3-dihydro-2,2-diméthyl-7-benzofuranyl	25,00	2	42,40	14,40	25,50
25	55285-14-8	FMC 35001	51,00	3	45,90	23,30	47,60
26	65907-30-4	Furathiocarb	53,00	3	43,90	26,30	35,80
27	82560-54-1	Benfuracarb	105,00	3	48,90	27,40	56,20
28	0-22-0	(2,2-Diméthyl-3H-1-benzofuran-7-yl) N-(4-bromophényl)sulfanyl-N-méthylcarbamate	25,00	2	37,20	25,00	N/A
29	0-70-6	(2,2-Diméthyl-3H-1-benzofuran-7-yl) N-(4-chlorophényl)sulfanyl-N-méthylcarbamate	25,00	2	38,10	22,30	N/A
30	7-88-8	Méthyl (1Z)-N-{{[[[2,2-Diméthyl-2,3-dihydro-1-benzofuran-7-yl]oxy]carbonyl](méthyl)amino]sulfanyl}(méthyl)carbamoyl]oxy)ethanimidothioate	15,90	2	40,10	18,10	15,70
31	8-14-9	(2,2-diméthyl-3H-1-benzofuran-7-yl) N-(4-tert-butylphényl)sulfanyl-N-méthylcarbamate	125,00	3	38,90	23,60	N/A
32	9-03-0	2,2-Diméthyl-2,3-dihydro-1-benzofuran-7-yl {{{(Z)-1,4-dithian-2-ylidèneamino]oxy]carbonyl}(méthyl)amino]sulfanyl}carbamate de méthyle	20,00	2	43,20	17,60	11,60
33	1563-67-3	Ester méthylique de l'acide carbamique, 2,3-dihydro-2-méthyl-7-benzofuranyl	43,00	2	8,78	2,65	48,70

2. Étude de cas 2

Tableau S2 : Les 29 substances chimiques membres du groupe des phosphorothiates sont référencées avec leur numéro CAS et leur nom de substance. Les valeurs expérimentales retrouvées, ainsi que les prédictions en read-across et QSAR sont indiquées pour chaque substance.

N°	CAS	Nom de la substance	Valeur expérimentale (mg/kg p.c.)	Classification CLP des données expérimentales	Valeur moyenne de lecture croisée (mg/kg pc)	Valeur minimale de lecture croisée (mg/kg pc)	QSAR (mg/kg pc)
1	39624-86-7	1,1'-Binaphthyl-2,2'-diylhydrogenphosphonate	12	2	47,80	9,27	20,70
2	2921-88-2	Iorsban	82	3	16,70	2,82	18,60
3	56-38-2	Parathion	2	1	12,70	3,00	17,00
4	97-17-6	Dichlofenthion	172	3	8,77	3,26	32,60
5	108-35-0	Pyrazothion	12	2	16,80	4,40	14,30
6	52-60-8	Acide phosphorothioïque O-[3,5-diméthyl-4-(méthylthio)phényl] O,O-diéthyl ester	375	4	17,60	4,44	49,50

7	297-97-2	Thionazin	3,5	1	9,93	4,40	55,30
8	298-03-3	Demeton-O	7,5	2	26,90	13,30	22,34
9	1716-09-2	Fenthion-éthyl	14	2	41,70	24,80	49,20
10	2425-15-2	Fénitrothion éthyle	9,99	2	58,70	22,00	25,30
11	3070-19-7	Chlorthion éthylrique	50	2	44,80	23,60	32,90
12	4824-78-6	Bromophos éthyl	51,9	3	38,90	23,60	46,10
13	4891-54-7	Déméton-o sulfone	89,9	3	32,40	22,30	42,29
14	5823-25-6	AI 3-23445	999	4	35,50	25,00	42,17
15	6012-87-9	Dicapthon éthylrique	10	2	80,00	31,50	45,32
16	7533-79-1	Acide phosphorothioïque, O-(2,5-dichloro-4-iodophényl) O,O-diéthyl ester	140	3	17,70	3,41	23,60
17	13457-18-6	Pyrazophos	218	3	18,00	3,59	25,10
18	18854-01-8	Isoxathion	112	3	27,80	25,70	94,68
19	21923-23-9	Chlorthiophos	13	2	45,00	22,30	30,30
20	29689-00-7	Acide phosphorothioïque, ester de O,O-diéthyle O-(2-nitrophényl)	25	2	50,00	21,00	110,00
21	33841-12-2	Acide phosphorothioïque, ester de O-(4-cyanophényl) O,O-diéthyle	10	2	39,30	24,40	26,80
22	42509-80-8	Miral	27	2	39,30	24,40	26,80
23	63873-23-4	Acide phosphorothioïque, ester de O,O-diéthyle O-(1-méthyléthyle)	1000	4	61,70	26,40	23,10
24	84197-35-3	Acide phosphorothioïque, O-(4-chloro-3-nitrophényl) O,O-diéthyl ester	100	3	42,40	14,40	25,50
25	84197-36-4	Acide phosphorothioïque, O-(2,4-dichloro-5-nitrophényl) O,O-diéthyl ester	100	3	45,90	23,30	47,60
26	1-57-8	O-(4-Chloro-2-nitrophényl) O,O-diéthyl phosphorothioate	1000	4	43,90	26,30	35,80
27	8-57-9	Mol-8580	10	2	48,90	27,40	56,20
28	14683-75-1	diéthoxy-hexane-2-yloxy-sulfanylidène-λ5-phosphane	350	4	37,20	25,00	N/A
29	20195-07-7	Acide phosphorothioïque, ester de S-butyle O,O-diéthyle	48,5	2	38,10	22,30	N/A

3. Étude de cas n° 3

Table S3 : Les 24 substances chimiques membres du groupe des esters organophosphorés sont référencées avec leur numéro CAS et leur nom de substance. Les valeurs expérimentales retrouvées, ainsi que les prédictions en read-across et QSAR sont indiquées pour chaque substance.

N°	CAS	Nom de la substance	Valeur d'observation	Valeur prédite (DL50 moyenne)	Valeur prédictive (DL50 minimale)	QSAR
1	78-30-8	PHOSPHATE D'O-TOLYLE	1160	4210	1450	3000
2	78-42-2	phosphate de tris(2-éthylhexyle)	37000	6960	1670	10000
3	115-86-6	butoxyéthyl)phosphate	3000	3710	1660	2800
4	126-71-6	disflamoll tp	3500	4930	1360	2700
5	126-73-8	Phosphate de triisobutyle	5000	2070	1110	1300
6	981-40-8	TRIBUTYLPHOSPHATE	3000	2290	1110	1700
7	29761-21-5	phosphate de t-butylphényldiphényle	1580	4400	1200	2400
8	57583-54-7	PHOSPHATE_D'ISODÉCYDE_DIPHÉNYLE	15800	7140	1610	1500
9	26444-49-5	Bis(phosphate de diphényle) de résorcinol	5000	2210	1030	550
10	132-29-6	Phosphate de diphényl crésyl	1420	6560	2520	2700
11	1806-54-8	Phosphate de 2-biphényle et de diphényle	1580	4680	1270	1600
12	2528-38-3	Phosphate de trioctyle	9550	8720	1670	
13	3862-08-6	Phosphate de tripentyle	23400	3330	1290	4700
14	36519-00-3	Phosphate de 2-éthylphénol (3:1)	15800	4670	1580	3300
15	52736-14-8	Phosdiphen	6200	3480	1310	1200
16	64532-94-1	phosphate de tris(3-éthylphényle)	15800	4670	1580	3300
17	64532-95-2	Phosphate de diphényle o-isopropylphénylphényle	10000	2740	1160	3000
18	64532-96-3	phosphate de tris(2-propan-2-ylphényle)	15800	8400	1740	3100
19	121-06-2	Phosphate de 4-hexylphényle diphényle	15800	4670	1580	3100
20	1323-22-4	Phosphate de tris(2,6-xylényle)	1580	7400	1580	2800
21	3862-12-2	Acide phosphorique, ester de di(2-éthylhexyle) hydroxypropyle	11300	5970	1590	7700
22	4281-67-8	Phosphate de tris(2,4-xylényle)	1580	7400	1580	2800
23	6163-73-1	Phosphate de cyclohexyle et de phényle ((C6H11O)(PhO)2PO) (7CI)	2460	5710	1390	3200
24		Éthanol, 2-méthoxy-, phosphate (3:1)	16900	2850	2010	1500

III. L'intelligence artificielle au service de la prédiction toxicologique

A. Introduction

La toxicologie vise à identifier et caractériser la relation qui existe entre les expositions à des agents et leurs conséquences toxiques, à déterminer les mécanismes d'actions et à faciliter l'évaluation des risques associés aux agents toxiques⁵². Comme précédemment évoqué, cette évaluation est traditionnellement réalisée *in vivo*, et demande du temps, du travail, des frais important et un grand nombre d'organismes vivants, ce qui ne s'est pas toujours révélé le plus efficace et éthique¹⁵⁸. Pour ces raisons, les modèles proposés par ordinateur représentent des alternatives prometteuses, impliquant une réduction des coûts, une suppression de la souffrance animale, une disponibilité rapide et à grande échelle des données ainsi que la standardisation des tests et procédures.

La toxicologie a subi de nombreuses transformations en quelques décennies, passant d'une science purement observationnelle à une discipline capable de prédiction *in silico* d'effets toxiques, comme c'est aujourd'hui le cas pour les candidats médicaments, les médicaments, les aliments, les produits chimiques¹⁵⁹.

La toxicologie prédictive repose sur l'intégration et l'analyse de données issues de sources variées, telles que les études sur les animaux, l'épidémiologie, les rapports cliniques, les essais *in vitro* et les approches « omiques »¹⁶⁰. Ce domaine de la toxicologie pourrait bénéficier de l'intégration de l'intelligence artificielle (IA), capable de répondre à la quantité et la complexité des données, fournissant des algorithmes capables d'apprentissage et d'analyse. L'intelligence artificielle fait référence à un ensemble de systèmes informatiques qui peuvent effectuer des tâches auparavant réservées au vivant, voire à l'intelligence humaine : la perception visuelle, la reconnaissance du langage, la prise de décision

et la traduction²². Forte de nombreuses avancées au cours de la dernière décennie, nous allons voir comment, appliquée à la toxicologie, elle peut contribuer à réduire les risques liés à l'exposition des individus, mais aussi à développer de nouveaux traitements et diagnostics des intoxications. Enfin, elle pourrait permettre une toxicologie personnalisée *via* des données génétiques et épigénétiques¹⁶⁰.

L'intelligence artificielle impacte des aspects clé en toxicologie comme :

- La toxicologie prédictive : les modèles d'apprentissage automatique (machine learning en anglais) sont construits à partir de bases de données chimiques existantes et des profils toxicologiques nécessaires à prédire la toxicité potentielle de nouveaux agents chimiques. Ils peuvent accélérer le criblage et réduire les tests sur les animaux¹⁸.
- L'analyse de données : ces méthodes peuvent permettre l'automatisation des activités comme l'exploration des études animales existantes, l'extraction d'informations de la littérature scientifique, l'analyse des données de criblage à haut débit et l'intégration de données « omiques »¹⁶¹.
- L'évaluation des risques toxicologiques : les modèles d'IA peuvent renseigner sur les probabilités de risques, en tenant compte de l'incertitude et de la variabilité, permettant ainsi une évaluation robuste¹⁶².
- La recherche sur les mécanismes chimiques : même si les modèles d'IA sont principalement des « boîtes noires », non explicables, les différentes avancées permettent d'en connaître plus sur les mécanismes d'action des molécules¹⁶³.

L'IA a pu évoluer avec la puissance des ordinateurs, la mise à disposition de larges bases de données et les avancées dans les algorithmes de « machine learning »¹⁶⁴. Ces technologies se développent dans un contexte de demandes croissantes pour remplacer les tests *in vivo* par ces modèles *in silico*, pour des

raisons économiques, de gain de temps, et, plus récemment, en raison de la fiabilité et de la robustesse des modèles d'IA qui s'améliorent. Les exemples les plus courants d'application de l'intelligence artificielle au service de la toxicologie concernent la recherche d'information et l'extraction de données. Les ordinateurs sont en passe de pouvoir traiter des millions d'articles scientifiques, et, si les tableaux et figures posent encore un problème dans l'analyse, cela devrait s'améliorer avec le temps¹⁶⁵. L'analyse d'images médicales, les données toxicologiques provenant d'anciennes études, les documents accumulés difficilement accessibles (ex : littérature grise), la littérature scientifique libre de droit, les bases données de tests (ToxCast, Tox21) et les plateformes « omiques » auront aussi un très grand impact sur la toxicologie¹⁶⁵.

C'est dans ce contexte d'une contribution croissante de l'IA que, grâce au traitement de données massives par différents types de modèles, la justesse en toxicologie prédictive, mais aussi en pharmacovigilance augmente. Cela interroge sur les défis éthiques et réglementaires accompagnant l'utilisation de l'IA dans ce contexte, qui devra être soumise à des recommandations pour une intégration efficace en toxicologie.

B. Fondements de l'intelligence artificielle en toxicologie

L'intelligence artificielle et le « machine learning », bien que basés sur les mêmes principes, possèdent deux définitions différentes :

- L'intelligence artificielle se réfère à la reproduction de l'intelligence humaine à travers des machines ou logiciels, dans le but d'exécuter des tâches qui requièrent une telle intelligence. On distingue l'intelligence artificielle spécifique (ANI ou « Artificial Narrow Intelligence »), appliquée dans un domaine précis comme l'assistant vocal ou la génération d'un itinéraire sur

un navigateur, de l'intelligence artificielle générale, capable d'effectuer une tâche cognitive propre aux humains¹⁶⁰.

- Le « machine learning » (aussi appelé apprentissage automatique) est une méthode d'analyse de données permettant l'application d'algorithme sans un programme explicite (figure 22). Les modèles apprennent à partir des données disponibles et s'améliorent de plus en plus en termes de justesse¹⁶⁰. Le machine learning inclut notamment la construction de modèles avec apprentissage supervisé, de classification et de régression (prédiction d'une variable continue). Les modèles de classification peuvent permettre de prédire l'occurrence d'un évènement, la mutagénicité par exemple, quand les modèles de régression servent à prédire des variables continues telles que la DL₅₀ (dose capable de tuer 50 % d'un groupe d'animaux d'essai)⁷⁷. De plus, il comprend l'apprentissage automatique non supervisé, qui identifie des groupes en fonction des données et par l'apprentissage par renforcement, souvent appliqué à des tâches de prise décision séquentielle, comme dans la conception de nouvelles molécules¹⁶⁶.

Ainsi, l'intelligence artificielle comprend de nombreuses approches dont le « machine learning », alors que ce dernier évoque spécifiquement des algorithmes basés sur l'apprentissage.

Il existe de nombreux systèmes d'intelligence artificielle, avec des types et des sous-domaines qui se chevauchent. Il existe de grands concepts clés tels que :

- Le Traitement automatique du Langage Naturel (« Natural Language Processing » ou NLP), qui se concentre sur la communication entre l'homme et l'ordinateur, fournissant aux machines la capacité de comprendre, interpréter et générer un langage humain de manière naturelle ;

- L'Apprentissage par Renforcement (« Reinforcement Learning » ou RL), qui permet d'adapter un agent autonome à un environnement, en lui faisant réaliser des actions qui amènent à une récompense (les renforcements). Il va se concentrer sur la maximisation de ces récompenses, ce qui va contribuer à la réalisation d'objectifs ;
- Les Réseaux de Neurones artificiels (« Artificial Neural Networks » ou ANN) qui sont des algorithmes utilisant les principes de l'organisation neuronale des cerveaux animaux et conçus pour reconnaître des motifs de données *via* la perception, l'étiquetage et le regroupement automatique ;
- Les Systèmes experts, qui reproduisent la prise de décision d'un expert humain afin de résoudre des problèmes complexes en utilisant le raisonnement et les connaissances spécialisées¹⁶⁰.

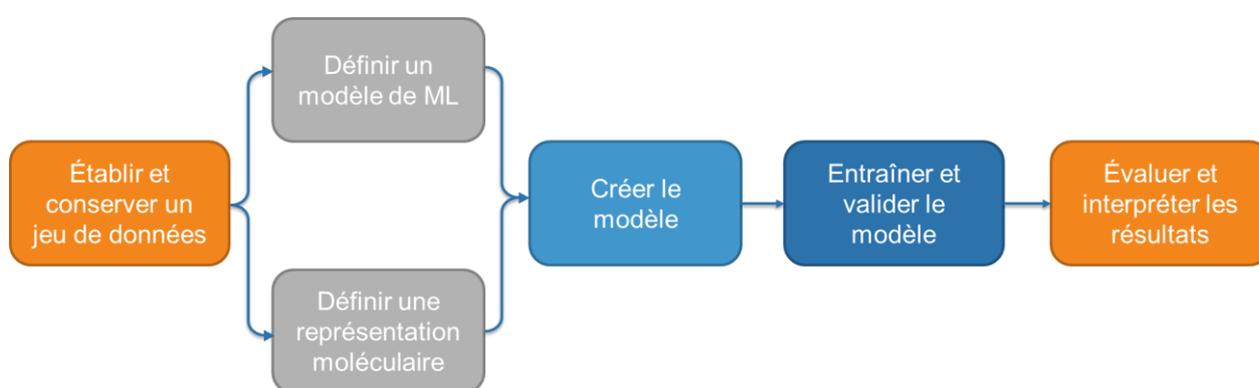


Figure 22 : Schéma de construction d'un modèle de « machine learning » (ML) (adapté de Cavasotto et al., 2022)¹⁶⁷.

Le machine learning et les approches utilisant l'intelligence artificielle se sont développées et sont de plus en plus utilisées pour répondre à des questions d'ordre toxicologique. Ces techniques sont aujourd'hui intégrées à des approches traditionnelles de prédiction et d'évaluation toxicologiques tels que les modèles QSAR (« Quantitative Structure-Activity Relationship »), les modèles PBPK (« Physiologically based pharmacokinetics ») ou encore l'analyse des AOP (« Adverse Outcome Pathways »)¹⁶⁸.

Les modèles QSAR, précédemment évoqués, représentent une méthode de simulation informatique pour étudier la relation quantitative entre les propriétés structurelles et les activités biologiques de substances chimiques. Largement utilisés en toxicologie prédictive et peu à peu intégrés à l'échelle réglementaire, l'arrivée de l'intelligence artificielle et du machine learning permet d'entraîner de nouveaux modèles et ainsi améliorer leur fiabilité. La méthode QSAR est basée sur des principes de machine learning traditionnels, et se distingue des méthodes dernier cri, très spécialisées et gourmandes en données. Certains considèrent les QSAR comme un type d'algorithme de machine learning¹⁵⁹.

Les modèles QSAR traditionnels, développés sans machine learning, bien que très souvent décrits dans la littérature, ont un pouvoir de prédiction limité. Ils reposent uniquement sur des descripteurs chimiques et ont une capacité réduite à couvrir de larges quantités de données, impactant ainsi leur utilisation dans toutes les classes chimiques¹⁶⁹.

A l'inverse, les algorithmes de machine learning ont le pouvoir de faire des prédictions complexes mais manquent parfois de clarté ou de compréhension par le commun des scientifiques ; ils possèdent des « boîtes noires », des composantes de fonctionnement non explicable¹⁵⁹.

La différence entre les deux concepts repose sur le fait qu'un modèle QSAR utilise la relation structure-activité pour permettre une prédiction quantitative d'un critère (équation linéaire), quand le machine learning traite les données disponibles par des techniques statistiques pour les généraliser. Ainsi, il permet d'obtenir des prédictions basées sur le modèle, utilisant ou non la structure moléculaire pour prédire le critère principal¹⁵⁹. Aujourd'hui, les algorithmes de machine learning sont utilisés pour construire des modèles QSAR capables de prédire l'activité d'un grand nombre de molécules¹⁶⁸, intégrant des données chimiques et biologiques diverses, qui permettent des prédictions basées sur les

données *in vitro*, *in vivo*, mais aussi les propriétés ADME (Absorption, Distribution, Métabolisation et Élimination)¹⁷⁰.

Les réseaux de neurones artificiels sont les plus utilisés aujourd'hui pour construire des QSAR¹⁷¹, en particulier les « Deep Neural Networks » (DNN ou réseaux de neurones profonds), un sous-type d'ANN mais possédant plusieurs couches, qui permet de traiter les données de manière complexe¹⁷². L'application de l'IA dans le domaine des QSAR a permis d'améliorer la justesse, les capacités d'extraction des données, les besoins en termes de puissance de calcul, l'interprétabilité, le transfert d'informations et l'optimisation des QSAR¹⁷³.

L'émergence de l'apprentissage profond (« deep learning » en anglais) au cours de la dernière décennie définit aujourd'hui ouvre la voie aux applications de l'IA en toxicologie, utilisant des réseaux de neurones multicouches pour extraire de nombreux caractères, à différents niveaux, provenant de données brutes¹⁸. De riches bases de données publiques telles que Tox21 et ToxCast enrichissent des méthodes de toxicologie prédictive telles que DeepTox¹⁷⁴, un modèle démontrant l'étendue de la puissance du deep learning dans la comparaison de substances chimiques.

Avec suffisamment de données pour entraîner les algorithmes, les modèles basés sur du deep learning peuvent saisir la complexité qui existe entre les structures chimiques, l'activité biologique et la toxicité. Une étude a notamment démontré toute la puissance du deep learning en l'adaptant à la prédiction de la mutagénicité : le caractère mutagène de plus de 90 % de substances chimiques d'un groupe d'essai a pu être prédit *in silico* avec la même justesse que les études *in vitro* initiales¹⁷⁴.

Les progrès de l'IA en toxicologie vont de pair avec ceux réalisés dans la disponibilité des données, les capacités des algorithmes et la collaboration entre différentes disciplines scientifiques. L'IA, notamment grâce au deep learning, offre

aujourd'hui la possibilité d'appréhender toutes les complexités de ces ensembles de données (« big data ») toxicologiques. Les bases de données publiques et les algorithmes construits en collaboration entre académiques et industriels, encouragent la transformation de la toxicologie prédictive et le développement de projets de recherche pour la diminution des effets indésirables liés aux substances chimiques et médicaments.

C. Applications de l'intelligence artificielle en pharmacologie et toxicologie prédictive

1. Prédiction de la cardiotoxicité par blocage du canal potassique (gène hERG)

La toxicité cardiaque est l'effet indésirable retrouvé le plus fréquemment dans le développement de nouveaux médicaments, conduisant à des troubles du rythme tels que l'arythmie¹⁷⁵. Un tel trouble est principalement dû à l'inhibition du canal potassique dépendant du gène hERG (« Human Ether-a-go-go ») par les médicaments cardiotoxiques.

Le gène hERG code pour la protéine Kv11.1, la sous-unité alpha du canal potassique cardiaque, nécessaire à la repolarisation du potentiel de membrane cardiaque¹⁷⁶. Ce canal dépendant du gène hERG est donc largement impliqué dans la bonne coordination du rythme cardiaque, qui peut être affecté au cours de certaines pathologies.

Une attention grandissante est portée sur les propriétés inhibitrices des composés bioactifs, leur identification précoce et de nombreuses études *in silico*, *in vitro* et *in vivo* ont été réalisées dans le but de prévenir d'éventuels effets cardiaques. L'objectif principal de ces études a été d'éclaircir les effets sous-jacents qui conduisent à ces effets cardiaques et comment les prévenir au cours du développement de nouveaux médicaments¹⁵⁸. Afin de prédire de potentiels effets, il a été nécessaire de procéder à une évaluation de l'activité envers le gène hERG,

après un criblage permettant de repérer les molécules « hits » (démontrant une activité biologique prometteuse, justifiant ainsi d'un examen approfondi) parmi les composés actifs.

Parmi ces études de détection précoces des effets anti-hERG, de nombreuses techniques de machine learning ont été utilisées, notamment celle de la « random forest » (RF), une méthode générant un grand nombre d'arbres de décision, un algorithme d'apprentissage supervisé très courant. Venkatraman (2021) a utilisé cette technique pour prédire les propriétés ADME des molécules candidates dans le développement pharmaceutique, en utilisant des modèles de classifications pour 50 critères ADME¹⁷⁷. Ces modèles de classification sont basés sur 20 empreintes moléculaires (systèmes de comparaison des caractéristiques chimiques des composés, afin d'en évaluer la similarité *in silico*¹⁷⁸), utilisés en routine dans le développement de nouveaux médicaments. Concernant le modèle de prédiction pour hERG, une justesse de 80 % a été obtenue, signifiant que 80 % des prédictions de classification étaient correctes. En utilisant le même modèle, d'autres critères ont été étudiés dans cette même étude (perméabilité de la barrière hémato-encéphalique, inhibition des CYP 450, liaison aux protéines plasmatiques, etc), parvenant au même taux de justesse. Avec les mêmes bases de données, Cai (2019) a pu développer un modèle de deep learning, « deepHERG », construit grâce à un réseau de neurones profonds multitâche (MT-DNN), surpassant les modèles simples de *machine learning* atteignant une justesse de 92 %¹⁷⁹. De plus, cet algorithme a mis en lumière l'activité d'inhibition potentielle de hERG par 29,6 % des médicaments déjà commercialisés, sur un échantillon de 1824 médicaments déjà approuvés par la FDA (Food and Drug Administration), soulignant l'importance de la détection précoce de la cardiotoxicité. Enfin, Zhang (2022) a développé un modèle de prédiction du blocage d'hERG, HergSPred, qui regroupe les moyennes de prédiction de différentes méthodes : le DNN, l'arbre de décision, la « random forest » (RF) et

s'appuyant sur l'« extreme gradient boosting » (XGB)¹⁸⁰. HergSPred a ainsi surpassé les capacités de prédiction de nombreux modèles différents, notamment celui de Cai : avec le même jeu de données de 3 485 composés, ce nouveau modèle a atteint une justesse de classification de 99,2 %.

Les méthodes basées sur l'intelligence artificielle, aussi bien machine learning que deep learning, sont de plus en plus utilisées pour leur capacité prédictive précise *in silico*. Au vu du nombre de données disponibles autour de l'inhibition du gène hERG, l'intérêt porté à la prédiction de la cardiotoxicité chez les candidats médicaments s'intensifie au sein de nombreuses équipes et les modèles de prédiction se multiplient. Les justesses augmentent, et cela est plus que primordial au vu du nombre toujours grandissant de candidats médicaments cardiotoxiques.

2. Prédiction des atteintes hépatiques d'origine médicamenteuse (DILI)

L'atteinte hépatique d'origine médicamenteuse (en anglais Drug-Induced Liver Injury, ou DILI) est associée à un surdosage en médicament, mais peut également survenir à dose thérapeutique. Ainsi, une toxicité directe découlant d'un surdosage en médicament est désignée sous le terme d'hépatotoxicité intrinsèque (réaction dose-dépendante), tandis qu'une hypersensibilité à un médicament est appelée hépatotoxicité idiosyncratique (non prévisible et beaucoup plus rare)¹⁸¹. Ces affections sont responsables du plus grand nombre de retraits de médicaments du marché¹⁸².

L'hépatotoxicité est le sujet de nombreux travaux de recherche pour accélérer le développement de nouveaux modèles de prédiction. Des approches par intelligence artificielle, le plus souvent basées sur le machine learning, permettent le développement d'outils capables de fournir des prédictions en analysant la structure chimique des médicaments mis en cause¹⁸³. Une connaissance plus précise des mécanismes sous-jacents aux DILI, mais également des motifs

moléculaires responsables des effets hépatotoxiques, permet le développement de modèles de prédictions reproductibles et fiables.

Dès 2006, O'Brien avait développé une approche qui classait plus de 600 composés par criblage haut débit sur des cellules humaines, les classant de non toxiques à sévèrement toxiques en comparant les résultats à des données obtenues grâce à des tests traditionnels¹⁸⁴.

Plus récemment, Chen (2016) a mis en place DILRank, un nouvel outil de classification des composés en fonction de la sévérité de la DILI induite, et dont plusieurs versions ont permis d'améliorer la recherche sur la responsabilité des médicaments dans l'induction des DILI¹⁸⁵. Capable de classer 1039 substances, cette base de données est la plus grande existante capable de prédire le risque de DILI. La taille des jeux de données ne cesse d'augmenter, notamment par la réunion de plusieurs d'entre eux¹⁸⁶.

De plus, des outils tels que Tox21, fruit d'une collaboration de plusieurs organismes américains, entre autres l'EPA et la FDA¹⁵⁷, permettent de développer des méthodes alternatives d'évaluation des risques des substances chimiques sur la santé humaine. En utilisant des techniques de criblage haut débit, Tox21 a pu fournir sur une bibliothèque de plus de 10 000 composés chimiques (dont la toxicité pour certaines voies biologiques est connue) avec plus de 100 millions de points de données¹⁸⁷. Cette base de données toxicologiques, disponible publiquement à l'intention de la communauté scientifique, permet d'améliorer les modèles de prédiction et les rendre exploitables pour les critères de sécurité.

Afin de développer des stratégies de prédiction des stades précoces de DILI, certaines études ont été menées pour classer les composés en fonction de leurs caractéristiques chimiques en lien avec une toxicité, permettant l'identification d'alertes structurelles : en plus de fournir des informations toxicologiques, elles renseignent sur les mécanismes sous-jacents conduisant à l'hépatotoxicité, ainsi

que sur le métabolisme des composés (réactivité chimique et biotransformation)¹⁸⁸.

En 2016, Pizzo et ses collaborateurs ont compilé des données pour 950 molécules de structure chimique connue qu'ils ont ensuite classées dans des groupes de squelette chimique similaire, mais avec des substituants différents. Cette approche a permis de formuler des hypothèses sur les alertes structurelles responsables d'hépatotoxicité. Ainsi, les alertes structurelles identifiées ont été utilisées pour créer un arbre décisionnel binaire permettant de prédire l'hépatotoxicité¹⁸⁹.

Les travaux de prédiction *in silico* des DILI sont grandissants, et de nombreux modèles sont mis en place. Cette accélération dans le développement de nouvelles techniques informatiques d'évaluation du risque hépatotoxique est nécessaire pour répondre aux problématiques du développement de nouvelles substances thérapeutiques, mais également pour répondre aux problématiques cliniques.

3. Prédiction de l'inhibition de l'acétylcholinestérase (AChE)

L'acétylcholinestérase est une enzyme responsable de la dégradation de l'acétylcholine, ainsi que d'autres neuromédiateurs possédant des fonctions esters de choline. Elle se trouve principalement dans les jonctions neuromusculaires et les synapses de transmission cholinergiques, et son inhibition est la cible des organophosphorés, utilisés comme insecticides et agents neurotoxiques (gaz de combat)¹⁹⁰. Elle est impliquée dans de nombreuses fonctions biologiques, notamment la croissance neurale, le développement osseux, la prolifération et l'apoptose des cellules hématopoïétiques¹⁹¹.

Certains inhibiteurs de l'acétylcholinestérase sont aussi utilisés en thérapeutique dans le traitement symptomatique de la maladie d'Alzheimer. Cette dernière est la première cause de démence chez la personne âgée, due à l'altération de la protéine bêta-amyloïde et hyperphosphorylation de la protéine tau¹⁹². Les symptômes incluent des signes précoces de pertes de mémoire, des troubles cognitifs et des difficultés d'expression. La cause de la maladie d'Alzheimer est peu connue, des facteurs de risques incluent des facteurs génétiques, l'âge, la dépression et l'hypertension¹⁹³.

Selon l'hypothèse dite « cholinergique », la maladie d'Alzheimer serait due à la réduction de la synthèse de l'acétylcholine et donc de la transmission de ce neurorécepteur au niveau du cerveau¹⁹⁴. Ainsi, le traitement à partir d'inhibiteurs de l'acétylcholinestérase est utile afin d'augmenter le niveau d'acétylcholine et ainsi la communication entre les cellules. Quatre molécules inhibitrices de l'acétylcholinestérase, visant les symptômes de la maladie d'Alzheimer (troubles cognitifs et de la mémoire) ont aujourd'hui une autorisation de mise sur le marché en France : le donepezil, la rivastigmine, la galanthamine et la tacrine. Cette dernière molécule est cependant rarement utilisée du fait de son hépatotoxicité¹⁹⁵.

Combinée à des modèles de classification, l'utilisation du machine learning pour la découverte de nouvelles molécules pour le traitement de la maladie d'Alzheimer est prometteuse. Sandhu, en 2022, a utilisé plusieurs modèles d'algorithmes combinés de machine learning (« support vector machine », « *k*-nearest neighbor » et « random forest ») pour distinguer les molécules inhibitrices l'acétylcholinestérase de celles qui ne le sont pas, et ainsi conduire à la sélection d'une large bibliothèque de molécules candidates. Les empreintes moléculaires pertinentes pour évaluer la robustesse des prédictions ont été choisies en amont, pour atteindre une justesse de 85,38% avec l'algorithme « random forest »¹⁹³.

Plus récemment, Khan (2023) a utilisé trois types d'algorithmes basés sur le machine learning pour identifier des molécules actives contre l'acétylcholinestérase. Parmi la base de données expérimentale (« Maybridge library », une base de données de 51 775 molécules), regroupant des valeurs d'IC50 de substances actives contre AchE, un petit groupe a retenu l'attention des chercheurs. Suite à la simulation des dynamiques moléculaires (MD) mais aussi l'analyse de l'affinité pour l'AchE, quatre composés semblent présenter des propriétés de liaison plus importantes que le Donepézil de référence. En effet, le score d'affinité PLP (« piecewise linear potential », qui calcule la complémentarité de la forme et la force de la liaison hydrogène entre le ligand et le site actif¹⁹⁶) du composé le plus puissant des quatre s'élevait à -123.99 contre -102.7 pour le Donepézil¹⁹⁷.

L'inhibition de l'acétylcholinestérase représente un enjeu pour le développement de nouvelles substances thérapeutiques contre la maladie d'Alzheimer. Malgré un nombre encore faible de médicaments actuellement sur le marché, les techniques d'intelligence artificielle, permettant du criblage à grande échelle de molécules avec prédiction de l'activité thérapeutique, devraient pouvoir accélérer le processus de développement.

D. Applications de l'intelligence artificielle en pharmacovigilance

La pharmacovigilance, science de la surveillance et de l'évaluation de la sécurité des médicaments, joue un rôle vital dans la compréhension et la prévention d'effets indésirables et inattendus liés à l'utilisation de médicaments¹⁹⁸. La sécurité qui entoure les nouveaux médicaments mis sur le marché est un défi majeur : en effet, malgré de nombreux tests rigoureux au cours d'essais cliniques, l'apparition d'une toxicité inattendue à la suite d'une mise sur le marché peut être responsable de morbidité ou mortalité inattendue. Le cadre d'un essai clinique est en effet éloigné de la réalité, comprenant un petit groupe

de participants, répondant à des critères précis : pas trop âgés, avec peu de comorbidité et sans polymédication, prenant le traitement étudié de façon contrôlée, pendant une durée limitée et sous étroite surveillance¹⁹⁹.

Le processus de pharmacovigilance s'assure donc de la détection et la surveillance des effets indésirables qui peuvent survenir dans des conditions réelles d'utilisation. Les données principalement utilisées en pharmacovigilance proviennent de déclarations individuelles de cas d'effets indésirables, qui sont généralement regroupées dans de larges bases de données. Il existe de nombreuses sources différentes, comme les dossiers médicaux électroniques, la littérature, les registres de patients, les discussions avec des chatbots, des programmes d'accompagnement de patients ou directement des réseaux sociaux²⁰⁰.

Le défi central de la pharmacovigilance est de pouvoir faire le tri parmi toutes ces sources d'informations et détecter les « signaux » (l'association causale entre un médicament et un effet indésirable) de manière rapide et précoce⁵¹. Ces nouvelles techniques permettent de considérer l'effet systémique d'un médicament, pour étudier les effets indésirables résultant d'interactions complexes sur l'organisme entier, plutôt que sur une seule cible ou voie de métabolisation. Alimenté par les grandes bases de données disponibles, le machine learning permet d'intégrer un grand nombre de caractéristiques à la prédiction *in silico*. Étant donné la progression de la place de l'IA dans de nombreux domaines scientifiques, elle devrait pouvoir résoudre les difficultés entourant une pratique optimale de la pharmacovigilance²⁰¹.

L'ANSM (l'Agence Nationale de la Sécurité du Médicament et des produits de santé) est l'autorité compétente en matière de pharmacovigilance au niveau national, prenant part à une organisation européenne pour la surveillance des effets indésirables des médicaments. Souhaitant intégrer pleinement l'intelligence artificielle à la pratique de la pharmacovigilance, elle lance en avril 2021 une

nouvelle application nationale qui permet une catégorisation et un précodage des effets indésirables, afin de faciliter et fluidifier le travail mené par les CRPV (Centres Régionaux de Pharmacovigilance). Cela a permis, durant la pandémie de Covid-19, l'analyse rapide des signalements d'effets indésirables rencontrés avec les médicaments et vaccins²⁰². Dans ce contexte de signalement, l'IA permet ainsi de faire face à la volumétrie importante des notifications spontanées arrivant aux CRPV, et ainsi centrer l'activité de ces derniers sur l'expertise médicale. Ainsi, l'IA contribue à automatiser les tâches et aider à la prise de décision, sans pour autant remplacer l'analyse et l'expertise médicales.

L'accès à de larges bases de données d'informations sur les patients se heurte à des difficultés techniques, notamment à cause de l'hétérogénéité de ces données mais aussi des problématiques légales et éthiques²⁰³. L'exploitation de bases de données rassemblant des dossiers médicaux électroniques (« electronic health records » ou EHR) permet de conserver des données démographiques, sur les traitements habituels, des résultats de tests de laboratoire ou encore des rapports de diagnostics cliniques²⁰⁴. Cela représente une possibilité de créer de nouveaux systèmes de classification des patients mais aussi de mettre en lumière des problématiques de pharmacovigilance. Afin d'identifier des signaux concernant la sécurité du médicament, présents dans les dossiers médicaux électroniques, différentes méthodes analytiques intégrant le « machine learning » peuvent être utilisées.

Le programme Sentinel, initié par la FDA en 2008, vise à développer des logiciels pour analyser des données de santé et évaluer la sécurité des médicaments déjà présents sur le marché²⁰⁵. Il utilise par exemple le traitement naturel du langage (« Natural Language Processing » ou NLP) pour extraire des informations utiles de ces quantités massives de données cliniques non structurées²⁰⁶. Le NLP permet de faire le lien entre l'informatique et la linguistique : en utilisant l'intelligence artificielle, il va créer des outils capables d'analyser, comprendre et générer un

langage humain²⁰⁷. Ainsi, l'extraction des données de santé grâce au NLP fournit une source pour le programme Sentinel pour la surveillance des médicaments et permet ainsi de générer des algorithmes d'identification des signaux de sécurité. Cependant, malgré l'intérêt porté à l'utilisation du dossier médical électronique, les travaux actuels ne parviennent pas à exploiter pleinement l'étendue des données disponibles, ni à contrôler rigoureusement les facteurs de confusion²⁰⁸.

L'intervention de l'IA en pharmacovigilance peut servir à soutenir une toxicologie fondée sur les preuves (EBT ou « evidence based toxicology »), une pratique qui permet d'analyser de façon critique des données de toxicité disponibles de manière transparente, rigoureuse et objective, pour résoudre un problème de sécurité²⁰⁹. En faisant un meilleur usage des ressources, la qualité des résultats et la facilité d'interprétation augmentent.

Afin de réaliser des revues systématiques de la littérature, le NLP peut permettre d'automatiser l'extraction d'informations médicales d'intérêt, parcourir les nombreux rapports d'effets indésirables et d'interactions médicamenteuses existants et les résumer, sans omettre de données capitales²¹⁰. Cette procédure effectuée en routine peut ainsi augmenter son efficacité : après un criblage préalable, cette méthode permet de trier les cas de pharmacovigilance présents dans des séries d'articles, du plus au moins urgent, et ainsi réduire le temps de prise en charge des cas rencontrés²¹¹.

Les interactions médicamenteuses sont l'une des principales causes d'effets indésirables évitables, pris en compte dans le cadre de la pharmacovigilance. Prédire de tels événements manuellement implique beaucoup de temps et de travail par des experts, mais aussi un manque d'efficacité couplé à l'erreur humaine. L'intérêt de l'intelligence artificielle est sa capacité d'analyse de bases de données étendues et complexes, mais aussi de cas qui pourraient passer inaperçus aux yeux des scientifiques.

Par exemple, Raja (2017) a utilisé des modèles de « random forest », issus du machine learning, pour prédire des interactions médicamenteuses avec des médicaments utilisés en dermatologie, en utilisant les données de la littérature²¹². Après sélection d'un glossaire de médicaments et produits chimiques à l'aide de trois banques de données, les interactions médicamenteuses ont été mises en évidence suite à l'étude des différents types d'effets indésirables, incluant les effets à un niveau moléculaire, liés à la pharmacocinétique de même que les interactions n'engendrant pas d'effet indésirable connu. La classification obtenue grâce à l'algorithme a obtenu un score de performance de 0,87, qui a augmenté avec l'introduction des données toxicogénomiques. Ainsi, l'étude des interactions médicamenteuses évolue avec l'émergence de l'intelligence artificielle, nourrie notamment par les données concernant l'information génétique mais aussi l'accès à des documents cliniques, plus ou moins structurés.

De nombreux projets sont en cours pour améliorer la pratique de la pharmacovigilance grâce à l'IA, notamment concernant la prédiction des effets indésirables liés aux médicaments. Par exemple, Wang (2021) a démontré comment un modèle de machine learning pouvait expliquer le lien de causalité entre la prise d'un médicament et l'apparition de l'effet indésirable. Cela pourrait permettre de répondre à des hypothèses de type « que se passerait-il si... ? », ou plus précisément, dans le cas de la pharmacovigilance : « que se passerait-il si un patient prenait un médicament A plutôt qu'un médicament B ? ». Appliqué à l'insuffisance hépatique aigüe liée aux analgésiques et la mortalité liée au tramadol, ce modèle, InferBERT, permet de récapituler les facteurs de risques connus pour ces effets indésirables et identifier de potentielles prédispositions chez un patient²¹³. Ainsi, une telle approche basée sur l'IA est prometteuse pour établir un lien de cause à effet entre les données observées dans la littérature et les cas cliniques, permettant de mieux comprendre la relation entre les prises de médicaments et les observations cliniques.

L'utilisation du « machine learning » se pose aujourd'hui comme une nécessité à l'amélioration de la pharmacovigilance et ainsi la prévention des effets indésirables engendrés par les médicaments. Pour autant, l'utilisation de ces outils performants ne vient pas sans interrogations quant à leur mise à disposition en routine. En effet, ces pratiques de surveillance de sécurité médicale ne sont possibles que grâce à la mise à disposition de données de santé privées, ce qui implique des enjeux éthiques et réglementaires.

E. Intelligence artificielle : enjeux éthiques et réglementaires

Les progrès des techniques d'intelligence artificielle appliquées à la santé sont rapides, et leur intégration est déjà en cours dans divers établissements de santé. Les domaines d'applications sont multiples, englobant aspects diagnostiques, cliniques et de gestion de service : médecine prédictive, médecine de précision, aide à la décision diagnostique et thérapeutique, construction de robots compagnons, chirurgie assistée par ordinateur ou encore prévention en population générale²¹⁴.

Ces applications offrent un large éventail d'avantages potentiels, que ce soit pour les organisations, les professionnels de santé ou les patients. Toutefois, ces nouvelles technologies dans le domaine de la santé ne sont pas sans risque. Leur avancée rapide et leur introduction suscitent un mélange de fascination, d'enthousiasme et d'espoir pour améliorer les soins et les diagnostics, mais aussi des inquiétudes, de la méfiance et de nombreuses questions. L'évolution constante des technologies d'IA, l'incertitude quant à leurs effets sur les personnes, les établissements de santé et la société en général, ainsi que l'absence de réglementation claire, nécessitent une réflexion approfondie sur les problèmes éthiques et réglementaires liés au développement des systèmes d'IA et le positionnement à adopter à leur égard.

Ce développement des technologies d'IA soulève en premier lieu des questionnements liés aux enjeux éthiques et l'exploitation de données massives en santé. En effet, la collecte et l'utilisation d'une grande quantité de données (aussi appelée « big data ») est nécessaire à l'entraînement d'algorithmes utilisés dans l'IA. De nombreux enjeux éthiques sont donc liés à la nature même de ces données, de leur collecte à leur archivage, en passant par leur entreposage, leur traitement, leur utilisation et leur partage (figure 23).

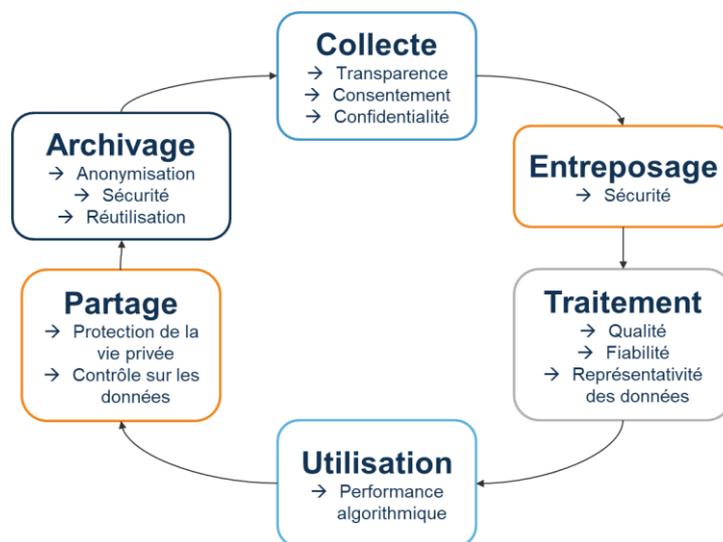


Figure 23 : Enjeux éthiques soulevés au cours du cycle de vie des données (adapté de Martineau et Gaudin, 2023)²¹⁵.

La collecte de données médicales se fait grâce à différentes sources, soit classiquement *via* les services de santé, la santé publique ou la recherche, soit de façon plus diverse *via* l'environnement (données de pollution, transport ou climat), le mode de vie et le statut socioéconomique (applications mobiles, cartes de fidélité, géolocalisation, transactions bancaires), sociales et comportementales (internet, réseaux sociaux, objets connectés qui fournissent des données de santé physique et de bien-être)²¹⁶. Ainsi, ces sources de données toujours plus importantes, alliées à de nombreuses possibilités de les étudier et la diversité des acteurs, participent au développement de nouvelles opportunités en matière de santé. Cependant, le partage des moyens de collecte entre les acteurs publics et privés compromet l'anticipation et la compréhension des utilisateurs qui se voient

soustraire beaucoup plus de données que nécessaires. De plus, l'utilisation commerciale ultérieure est souvent inconnue des individus auprès de qui ces données sont collectées²¹⁷. Le consentement à la collecte et au partage des données est quant à lui de plus en plus réglementé, mais des questions se posent quant à son obtention. Dans les faits, l'utilisateur est souvent exposé à un jargon juridique inaccessible, manquant d'explication claire, qui l'empêche de comprendre les implications de ce consentement²¹⁷. Ce manque de transparence interroge donc sur la pleine autonomie des personnes dans la collecte et le partage de leurs données de santé.

La confidentialité et la vie privée des individus concernés est également sujette à des questionnements éthiques. Pour assurer la protection de ces informations sensibles, les différents acteurs utilisent des méthodes pour empêcher l'identification des individus. Or, cette confidentialité peut être mise à mal par diverses failles ; une étude a notamment démontré que, si des individus sont régulièrement localisés sur une carte, seulement quatre points de données spatio-temporels sont nécessaires pour identifier 95 % d'entre eux²¹⁸. Ainsi, des données censées être anonymisées (c'est-à-dire retirées des bases d'identifications de manière irréversible) pourraient être retrouvées par un acteur spécifique et permettraient de réidentifier les individus²¹⁹. La circulation massive de données augmente aussi le risque de cyberattaques, pouvant notamment paralyser les systèmes d'information hospitaliers (SIH), provoquant la paralysie de l'établissement de santé et la fuite de données sensibles²²⁰.

La qualité et la fiabilité des données collectées sont primordiales afin de ne pas biaiser les algorithmes développés sur ces données. Elles varient en fonction des différentes sources, la gestion de ces données, leur représentativité et leur quantité. Les données doivent être adaptées aux objectifs des algorithmes, car une incohérence dans l'application des critères garantissant la qualité, la fiabilité et la représentativité des données peut impacter la performance des algorithmes,

compromettant la qualité de leurs prédictions²¹⁹. La validation est ensuite nécessaire à la vérification de la qualité de performance de l'algorithme, et doit être réalisée avec un ensemble de données différent de celui utilisé pour son développement. En effet, il est nécessaire d'évaluer sa performance en cas d'exposition à un autre jeu de données pour éviter le sur-apprentissage. L'entraînement d'algorithmes sur une base de genre, ethnicité, âge ou positionnement peut en effet conduire à la discrimination d'une tranche de la population²¹⁵.

La dernière étape du traitement des données nécessaire à l'implémentation de l'IA dans la santé concerne l'archivage et la conservation des données. Cette dernière se fait souvent de façon anonymisée, même si le risque d'une réidentification est toujours existant, comme évoqué précédemment. De plus, cette perte d'identité s'accompagne de celle des mesures juridiques qui encadrent les conditions d'utilisation. Ainsi, elles pourraient servir à des fins différentes de celles qui ont été précédemment consenties²²¹. La perte de contrôle sur ces données peut également poser un problème aux individus vivant avec des conditions de santé particulières, qui peuvent ainsi se voir stigmatisés par le « profil digital » créé par le partage de leurs données, représentant ce que les différents acteurs du milieu de la santé connaissent sur eux²²².

L'intégration de l'IA dans le secteur de la santé offre des avantages, mais les différents acteurs impliqués dans la décision doivent évaluer attentivement l'éthique de ses applications. Certains systèmes peuvent enfreindre des principes éthiques, notamment à cause de la dissimulation des intentions des développeurs dans les conditions d'utilisation. Une montre intelligence ou une application d'enregistrement d'activités peut par exemple compromettre l'autonomie des utilisateurs si ces dispositifs sont utilisés à des fins de surveillance et de manipulation comportementale. L'évaluation du caractère éthique doit considérer la balance entre les coûts et les avantages, incluant les risques de discrimination.

Les organisations en charge se doivent de mettre en place des processus transparents pour anticiper, atténuer les risques et réagir rapidement en cas de préjudice causé par ces technologies.

À l'heure actuelle, les systèmes de décision algorithmique en santé agissent principalement comme des outils de soutien à la prise de décision humaine, n'ayant pas une autonomie décisionnelle complète. Bien que ces systèmes génèrent prédictions et recommandations, la complexité de certains systèmes d'IA crée une opacité dans leur fonctionnement, connue sous le nom de « boîte noire », entravant la compréhension, l'interprétabilité et l'acceptabilité par les professionnels de santé et les patients. L'incapacité à expliquer le raisonnement de l'IA peut entraîner des désaccords entre les cliniciens et les recommandations du système, créant des doutes sur la validité des décisions prises et entravant son intégration dans le système de santé. De plus, le jugement professionnel pourrait se voir atteint par une dépendance excessive à ces technologies.

Des erreurs potentielles dans les recommandations des algorithmes peuvent avoir des conséquences graves sur la qualité des soins et la santé des patients. Il est donc recommandé pour les entreprises développant ces systèmes d'améliorer leur explicabilité (c'est-à-dire leur capacité à rendre compréhensibles les éléments pris en considération par le système d'IA afin de produire un résultat²²³), afin de garantir une utilisation plus transparente et fiable de cette technologie en santé.

De plus, la complexité des décisions cliniques assistées par l'IA soulève des questions sur la responsabilité. En effet, si le clinicien n'a ni la compréhension ni le contrôle sur l'algorithme, peut-il être tenu moralement ou légalement responsable d'une erreur grave sur un patient ? La répartition de la responsabilité entre les acteurs impliqués dans le développement, la commercialisation et le déploiement des systèmes d'IA reste à définir. À cela s'ajoutent les biais créés lors de l'apprentissage des algorithmes, qui se fait en fonction des décisions humaines prises lors du développement. Les paramètres sociaux qui y sont intégrés,

transmis par les concepteurs, sont ainsi amplifiés et peuvent conduire à la discrimination d'individus appartenant à certaines communautés marginalisées²²⁴.

Ces systèmes autonomes capables de prise de décision sont également des agents moraux artificiels. Leur implication dans la prise de décision médicale soulève des préoccupations quant à la possible altération de la participation du patient dans le processus décisionnel. Les valeurs morales implicitement véhiculées par ces systèmes sont influencées par les biais et les intérêts individuels des développeurs, ce qui souligne la nécessité d'une réglementation et d'une supervision éthiques, plus poussées dans le développement des systèmes d'IA.

Le marché du travail sera aussi bouleversé par l'intégration de l'IA dans le milieu de la santé. Par souci d'autonomie et de dignité des travailleurs de la santé, il est nécessaire que ces systèmes capables de décisions ne dévalorisent pas l'expertise humaine en ne remplaçant pas la totalité de ses fonctions²¹⁹.

L'intégration de l'IA à la médecine, et plus précisément à la pratique de la toxicologie, présente à la fois des perspectives passionnantes et des enjeux éthiques et réglementaires considérables. Les avantages potentiels de l'IA doivent être suivis de près par toutes les parties prenantes, scientifiques comme gouvernementales, pour l'élaboration de mesures adaptées et applicables à l'échelle planétaire. Si les professionnels de santé ne sont pas destinés à être remplacés par des systèmes d'IA, ils devront néanmoins apprendre à coexister avec eux.

Plus spécifiquement appliqués à la toxicologie, Stephan Platz (AstraZeneca) énonçait au cours du congrès annuel de la British Society of Toxicology (2019) que « l'IA ne remplacera pas les toxicologues, mais les toxicologues qui n'utilisent pas

l'IA seront remplacés par ceux qui le font ». Les perspectives futures sont nombreuses dans ce domaine florissant, et l'IA peut transformer la toxicologie prédictive si elle reste conçue et encadrée par l'humain¹⁶⁰.

F. Défis, limites et perspectives futures

Le « Principe d'Anna Karénine » est basé sur la phrase d'ouverture du roman éponyme de Léon Tolstoï : « Toutes les familles heureuses le sont de la même manière, les familles malheureuses le sont chacune à leur façon ». Cela signifie que pour réussir à répondre à un problème complexe, il est nécessaire d'éviter tout obstacle à sa résolution²²⁵. Un seul élément manquant peut faire échouer le système. Dr Andreas Bender, Professeur de bio-informatique à l'Université de Cambridge, a adapté cette citation, changeant librement le terme de « famille » par celui de « molécules » : « all happy molecules are alike; each unhappy molecule is unhappy in its own way »²²⁶. Sachant que chaque molécule doit être efficace (présenter un certain nombre de propriétés) mais aussi doit être fiable en termes de sécurité (absence d'un certain nombre de propriétés toxiques), il exprime son opinion sur l'usage de l'IA dans la prédiction toxicologie : comment prédire un mode d'action toxique que l'on n'a jamais vu auparavant ? Ou bien la potentielle toxicité d'une nouvelle structure moléculaire ?²²⁶. La toxicité d'une molécule n'est pas nécessairement liée à son mécanisme d'action, ce qui la rend beaucoup difficile à anticiper.

Des limites se posent donc, notamment concernant :

- Les biais et la qualité des données : les modèles d'IA sont susceptibles de renforcer les biais présents dans les bases de données servant à façonner les algorithmes, pouvant conduire à des inexactitudes dans les prédictions. Il faut donc bien sélectionner les données, afin qu'elles soient représentatives, ne présentant pas de biais initiaux²²⁷. L'IA apparaît

aujourd'hui comme une technique de choix pour transformer l'approche probabiliste en évaluation du risque, c'est-à-dire une méthode quantitative qui définit la probabilité qu'un risque a de se produire. La variabilité et les incertitudes des données sont intégrées dans l'évaluation du risque pour sa caractérisation complète²²⁸. L'avantage de cette approche est de pouvoir accompagner les prédictions des incertitudes qui en découlent, afin d'affiner le modèle ultérieurement. Des techniques d'IA telles que le deep learning bayésien permettent d'explicitier les incertitudes dans les résultats des réseaux de neurones²²⁹, permettant d'obtenir des intervalles de confiance et des probabilités pour l'estimation des risques.

- Le manque d'uniformité : les formats et dénomination de données différents, mais aussi des protocoles de collecte et d'archivage hétérogènes peuvent rendre l'intégration de l'IA complexe. Il est nécessaire de mettre en place des lignes directrices et s'inscrire dans les principes FAIR (Findable, Accessible, Interoperable and Reusable) : des données facilement accessibles, comprises, échangeables et réutilisables²³⁰.
- L'expertise multidisciplinaire : le développement et l'implémentation des algorithmes d'IA nécessitent une collaboration entre experts en données (« data scientists »), ingénieurs informatiques, chimistes, biologistes et toxicologues²²⁸. Le projet européen ONTOX, créé en 2020, est l'illustration de cette volonté de collaboration pour intégrer l'IA à la toxicologie du XXI^e siècle²³¹. Ce projet vise spécifiquement à développer six NAM pour évaluer le risque *in silico* de la toxicité chronique, et la collaboration entre acteurs industriels et réglementaires devrait favoriser l'acceptation de cette évaluation de la sécurité chimique limitant l'impact animal.
- L'interprétabilité des modèles : les avancées dans le domaine de l'IA doivent pouvoir être expliquées, afin de comprendre le mécanisme sous-jacent des algorithmes. Certaines ressources telles que « Integrated Chemical

Environnement » sont construites pour une meilleure compréhension et un accès facilité aux différents outils d'IA à la communauté scientifique. Organisée par critère d'évaluation toxicologique, la base de données ICE facilite l'exploration des données publiques, des différents tests réalisés et des résultats observés au niveau biologique. L'utilisateur peut également simuler des essais *in vivo* virtuels, et l'intelligence artificielle lui permet d'en définir tous les paramètres (espèces, dosage, temps...) ²²⁸.

- L'acceptabilité réglementaire : afin de permettre l'approbation au niveau réglementaire, le développement du modèle et les hypothèses émises grâce à ce dernier doivent faire preuve de transparence. Il est important de vérifier, valider et quantifier l'incertitude des résultats. Le Comité de coordination inter-agences pour la validation des méthodes alternative (ICCVAM), aux États-Unis, a récemment publié un rapport afin de fournir un cadre pour la validation des nouvelles méthodologies ²³², de même que l'OCDE continue ses efforts pour l'évaluation réglementaire des méthodes QSAR, à travers le « QSAR Assessment Framework » ²³³.

Malgré ces limites, les opportunités promises par l'IA sont nombreuses. Tout d'abord, les modèles prédictifs *in silico*, se servant de données préexistantes, peuvent permettre de réduire l'utilisation animale pour l'évaluation toxicologique de nouvelles substances chimiques. De plus, l'IA permettrait d'automatiser certaines tâches fastidieuses, laissant plus de temps aux experts toxicologues pour se focaliser sur des tâches plus complexes, le tout encourageant une évaluation plus rapide de la sécurité. Enfin, les systèmes d'IA peuvent faciliter l'accès aux outils prédictifs et aux bases de données toxicologiques aux différents utilisateurs ²²⁸.

La mise en place de l'IA en toxicologie est encore à un stade précoce. Pour développer l'IA en recherche et en dans le domaine réglementaire de façon responsable et exploiter pleinement son potentiel, il est nécessaire de traiter les

questions d'éthique et de transparence, l'utilisation des données et l'interprétabilité des modèles. À mesure que les techniques d'IA évoluent, la confiance en ces systèmes gagne du terrain et ils devraient constituer un élément essentiel de la toxicologie, avec un champ d'applications s'étendant de l'étude des effets indésirables aux effets de variation d'une population, en passant par les effets « off-target », les biomarqueurs, les différences inter espèces et les AOP²²⁸.

Même si l'IA peut améliorer les capacités prédictives et l'automatisation, la créativité, l'esprit critique et l'expertise des chercheurs en toxicologie resteront indispensables pour utiliser ces outils novateurs en toute sécurité. Malgré son potentiel, l'IA entraîne questionnements et défis concernant les préoccupations éthiques et la protection des données, qui devront être traitées avec le plus grand intérêt. Une collaboration scientifique sera nécessaire pour surmonter les différences en termes de connaissances entre spécialistes de l'IA et chercheurs en science de la vie.

Conclusion

Il existe aujourd'hui un consensus entre la communauté scientifique et les autorités réglementaires sur l'importance de l'utilisation des NAM dans la toxicologie prédictive. Cette importance grandissante offerte aux méthodes alternatives est une opportunité de participer à l'évaluation du danger tout en étant plus protecteur envers la santé humaine mais aussi en respectant le principe des 3R – remplacer, réduire et raffiner l'expérimentation animale. Les incertitudes posées par le système d'évaluation actuel reposant en grande partie sur les tests animaux, les questions sur la pertinence des modèles utilisés, les substances chimiques trop peu testées et encadrées par le règlement REACH²³⁴, ou encore les nouveaux défis posés par les molécules candidates à la commercialisation, sont autant de problématiques qui peuvent être adressées par des approches basées sur les NAM.

Afin d'évaluer la pertinence de ces approches dans un contexte particulier, une étude de cas a pu être mise en place grâce à deux approches NAM *in silico*. Trois groupes de substances (33, 29 et 24 molécules), avec des similitudes structurelles et mécanistiques établies, ont été évaluées pour leur toxicité orale aiguë, en comparant les prédictions faites par read-across et QSAR à des résultats obtenus *in vivo*.

Le règlement CLP, permettant de classer les substances en fonction de leur toxicité aiguë, a permis d'évaluer la justesse des prédictions obtenues *in silico* : certaines ont été sous-estimées, principalement pour les composés les plus toxiques, et d'autres ont été surestimées, en particulier avec le read-across. Ainsi, intégrés dans un système de classification CLP, au sein d'un contexte précis, les résultats de ces NAM nécessitent d'être considérés, en particulier pour les composés à faible toxicité.

Ce travail est un aperçu de la nécessité d'impliquer d'autres études pour améliorer la justesse de prédiction et accroître la confiance dans les résultats des NAM, en analysant davantage de groupes de substances, présentant des variables supplémentaires susceptibles d'influencer les résultats. Les progrès réalisés dans le domaine de l'IA devraient permettre d'étendre les possibilités de méthodes envisagées pour l'évaluation de la toxicité des substances. En effet, ces progrès sont étroitement liés à l'augmentation de la disponibilité des données, aux capacités des algorithmes et à la collaboration entre différentes disciplines scientifiques. L'IA, en particulier le *deep learning*, offre la possibilité d'appréhender la complexité des ensembles de données toxiques, et favorise des applications en toxicologie prédictive incluant la prédiction de la toxicité potentielle de nouveaux agents chimiques, l'analyse d'images médicales, l'extraction de données à partir de documents accumulés difficilement accessibles, et l'interprétation des modèles.

Cependant, malgré son potentiel, l'IA soulève des questions éthiques et des défis en matière d'accessibilité et d'interprétabilité des modèles : elle se doit de faire ses preuves auprès des autorités réglementaires pour pleinement prendre part à la toxicologie prédictive.

Historiquement, la toxicologie réglementaire est sujette à des exigences légales rigides, étant directement impliquée dans la protection de la santé du grand public. La rigueur scientifique, la fiabilité des données mais aussi la nécessité d'une administration solide, composée de procédures prévisibles, ont produit des cadres hautement normalisés. Le système actuel est ainsi difficilement ouvert au changement, favorisant une sécurité optimale mais opposant des résistances à l'intégration de nouveaux essais, qui pourraient être adaptés à l'état de la recherche scientifique. Les NAM abordant souvent plusieurs aspects d'un critère d'évaluation, elles analysent les informations sur plusieurs niveaux d'ordre biologique et ne sont pas toujours restrictives au seul système

physiologique humain : elles ne sont pas toujours adaptées aux normes réglementaires établies.

Considérer les NAM comme fournissant des informations de nature différente, aussi adéquates voire plus pertinentes pour l'évaluation des dangers que les méthodes *in vivo*, pourrait favoriser leur intégration à un cadre réglementaire. Le cadre établi aborde les critères d'évaluation toxicologiques selon une approche basée sur l'expérimentation animale, et ne prend pas en compte toutes les variables étudiées grâce aux NAM. La législation actuelle devra donc passer à des méthodes probabilistes, pertinentes pour la santé humaine et bien informées au niveau mécanistique. Il est ainsi nécessaire d'évaluer et de comparer les performances des anciennes et nouvelles méthodes, ainsi que de réexaminer voire reformuler les exigences des tests existantes. Le concept de critère d'évaluation doit être élargi pour ne plus penser qu'en terme de point final mais en intégrant les évènements cellulaires et moléculaires sous-jacents. Les NAM existantes doivent être utilisées en réseau pour développer et optimiser des AOP, et ainsi aborder l'évaluation toxicologique de façon plus globale.

Chaque système comporte ses limites ; comparer la performance des systèmes traditionnels à celle des systèmes alternatifs risque de favoriser l'amplification de biais dans les évaluations. Ainsi, il est nécessaire que tous les acteurs impliqués – scientifiques, industriels et autorités réglementaires – collaborent pour l'adoption des NAM dans les cadres juridiques. La maîtrise des outils *in silico* disponibles s'accompagnera d'une plus ample connaissance de leurs bénéfices et risques, et conduira vers de nouvelles pistes de recherche entourant l'évaluation du danger en toxicologie.

Bibliographie

1. Roudot, A.-C. *Modélisation En Évaluation Du Risque Toxicologique*. (2022).
2. Caudeville, J. Évaluation et gestion des risques de mélanges de produits chimiques : défis et perspectives. *Environ. Risques Santé* **18**, 451–454 (2019).
3. Carrier, G. & Bard, D. Analyse du risque toxicologique. *Environ. Santé Publique - Fond. Prat.* 203–226 (2003).
4. Glorennec, P., Bonvallot, N., Noisel, N., Rousselle, C. & Jailler, M. Chapitre 14. Évaluation des risques sanitaires. in *Environnement et santé publique* 361–399 (Presses de l'EHESP, Rennes, 2023). doi:10.3917/ehesp.goupi.2023.01.0361.
5. Barlow, S. M. *et al.* Hazard identification by methods of animal-based toxicology. *Food Chem. Toxicol.* **40**, 145–191 (2002).
6. Mendes, H. Claude Bernard et l'origine de l'expérimentation animale moderne. *Bibnum Textes Fond. Sci.* (2012) doi:10.4000/bibnum.532.
7. Noisel, N., Bonvallot, N. & Coumoul, X. Chapitre 11. Toxicologie. in *Environnement et santé publique* 277–304 (Presses de l'EHESP, Rennes, 2023). doi:10.3917/ehesp.goupi.2023.01.0277.
8. Ineris. Méthodes alternatives à l'expérimentation animale : État des lieux et enjeux pour l'Ineris. (2023).
9. Chinwalla, A. T. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
10. Perlman, R. L. Mouse models of human disease: An evolutionary perspective. *Evol. Med. Public Health* **2016**, 170–176 (2016).
11. Hernandez, R. E., Galitan, L., Cameron, J., Goodwin, N. & Ramakrishnan, L. Delay of Initial Feeding of Zebrafish Larvae Until 8 Days Postfertilization Has No Impact on Survival or Growth Through the Juvenile Stage. *Zebrafish* **15**, 515–518 (2018).
12. Horzmann, K. A. & Freeman, J. L. Making Waves: New Developments in Toxicology With the Zebrafish. *Toxicol. Sci. Off. J. Soc. Toxicol.* **163**, 5–12 (2018).
13. Ministère de l'enseignement supérieur et de la recherche. Enquête statistique sur l'utilisation des animaux à des fins scientifiques. *enseignementsup-recherche.gouv.fr* <https://www.enseignementsup-recherche.gouv.fr/fr/enquete-statistique-sur-l-utilisation-des-animaux-des-fins-scientifiques-46270> (2024).
14. Parlement Européen. Produits cosmétiques testés sur les animaux: Vers une interdiction mondiale! *Thèmes | Parlement européen* <https://www.europarl.europa.eu/topics/fr/article/20180216STO98005/produits-cosmetiques-testes-sur-les-animaux-vers-une-interdiction-mondiale> (2018).
15. Hubert, P. & Toulhoat, P. Les méthodes alternatives en matière d'expérimentation animale. *Ann. Mines - Responsab. Environ.* **71**, 19–23 (2013).
16. Lysenko, A., Sharma, A., Boroevich, K. A. & Tsunoda, T. An integrative machine learning approach for prediction of toxicity-related drug safety. *Life Sci. Alliance* **1**, (2018).
17. European Chemicals Agency (EU body or agency). *The Use of Alternatives to Testing on Animals for the REACH Regulation*. (Publications Office of the European Union, LU, 2020).
18. Luechtefeld, T., Marsh, D., Rowlands, C. & Hartung, T. Machine Learning of Toxicological Big Data Enables Read-Across Structure Activity Relationships (RASAR) Outperforming Animal Test Reproducibility. *Toxicol. Sci.* **165**, 198–212 (2018).
19. CNRS. REACH | Data on substances > Information requirements for registration. https://www.prc.cnrs.fr/reach/en/data_requirements.html (2007).
20. Schmeisser, S. *et al.* New approach methodologies in human regulatory toxicology – Not if, but how and when! *Environ. Int.* **178**, 108082 (2023).

21. Cash, R. Chapitre 3. Les limites de l'expérimentation animale. *Essais* 105–131 (2022).
22. Pérez Santín, E. *et al.* Toxicity prediction based on artificial intelligence: A multidisciplinary overview. *WIREs Comput. Mol. Sci.* **11**, e1516 (2021).
23. Van Norman, G. A. Limitations of Animal Studies for Predicting Toxicity in Clinical Trials: Is it Time to Rethink Our Current Approach? *JACC Basic Transl. Sci.* **4**, 845–854 (2019).
24. Hartung, T. Toxicology for the twenty-first century. *Nature* **460**, 208–212 (2009).
25. Lonjon, N. *et al.* Minimum information about animal experiments: supplier is also important. *J. Neurosci. Res.* **87**, 403–407 (2009).
26. Pound, P. & Bracken, M. B. Is animal research sufficiently evidence based to be a cornerstone of biomedical research? *BMJ* **348**, g3387 (2014).
27. Leist, M. *et al.* Test systems of developmental toxicity: state-of-the art and future perspectives. *Arch. Toxicol.* **87**, 2037–2042 (2013).
28. Marguénaud, J.-P. L'expérimentation animale, entre droit et liberté. *Sci. En Quest.* 7–53 (2011).
29. Russell, W. M. S. & Burch, R. L. The Principles of Humane Experimental Technique. *Med. J. Aust.* **1**, 500–500 (1960).
30. Inserm. Qu'est-ce que la règle des 3 R ? · Inserm, La science pour la santé. *Inserm* <https://www.inserm.fr/modeles-animaux/qu-est-regle-3-r/> (2017).
31. Veissier, I. & Deiss, V. Éthique et réglementation en expérimentation animale en France: éléments de réflexion à partir des expérimentations sur les bovins. *INRAE Prod. Anim.* **35**, 155–170 (2022).
32. Dangles-Marie, V., Guichet, J.-L. & Larue, L. Éthique et expérimentation animale. in *Traité de bioéthique* 370–381 (Érès, Toulouse, 2010). doi:10.3917/eres.hirsc.2010.01.0370.
33. Poole, T. Happy animals make good science. *Lab. Anim.* **31**, 116–124 (1997).
34. Parlement Européen. *Directive 2010/63/UE du Parlement européen et du Conseil du 22 septembre 2010 relative à la protection des animaux utilisés à des fins scientifiques Texte présentant de l'intérêt pour l'EEE.* *OJ L* vol. 276 (2010).
35. Ministère de l'enseignement supérieur et de la recherche. Cadre réglementaire. *enseignementsup-recherche.gouv.fr* <https://www.enseignementsup-recherche.gouv.fr/fr/cadre-reglementaire-87211> (2022).
36. European Commission. *REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL 2019 Report on the Statistics on the Use of Animals for Scientific Purposes in the Member States of the European Union in 2015-2017.* (2020).
37. European Commission. European Chemical Strategy for Sustainability. (2020).
38. Balls, M. Replacement of animal procedures: alternatives in research, education and testing. *Lab. Anim.* **28**, 193–211 (1994).
39. Canada, S. Utilisation des nouvelles approches méthodologiques d'évaluation des risques. <https://www.canada.ca/fr/sante-canada/services/substances-chimiques/fiches-renseignements/utilisation-nouvelles-approches-methodologiques-evaluation-risques.html> (2023).
40. Fischer, I., Milton, C. & Wallace, H. Toxicity testing is evolving! *Toxicol. Res.* **9**, 67–80 (2020).
41. Goodman, J. I. Goodbye to the bioassay. *Toxicol. Res.* **7**, 558–564 (2018).
42. Smirnova, L. *et al.* 3S - Systematic, systemic, and systems biology and toxicology. *ALTEX - Altern. Anim. Exp.* **35**, 139–162 (2018).
43. Browne, P. *et al.* Development of a curated Hershberger database. *Reprod. Toxicol.* **81**, 259–271 (2018).
44. Delrue, N. & Francopa, S. NAM, les outils développés au sein de l'OCDE et orientations futures. (2022).
45. Cattaneo, I. *et al.* Implementing New Approach Methodologies (NAMs) in food safety assessments: Strategic objectives and actions

- taken by the European Food Safety Authority. *Trends Food Sci. Technol.* **133**, 277–290 (2023).
46. ECHA. New approach methodologies in regulatory science - Publications Office of the EU. <https://op.europa.eu/en/publication-detail/-/publication/4c2ad7eb-9ad0-11e6-868c-01aa75ed71a1/language-en> (2016).
47. ECHA. Grouping of substances and read-across - ECHA. <https://echa.europa.eu/support/registration/how-to-avoid-unnecessary-testing-on-animals/grouping-of-substances-and-read-across>.
48. ANSES. Évaluation du poids des preuves à l'Anses: revue critique de la littérature et recommandations à l'étape d'identification des dangers. (2016).
49. Mombelli, E. Évaluation de la reproductibilité des approches « read across » entre évaluateurs. *Rapp. Sci. INERIS* **2016–2017**, 62 (2017).
50. Ineris. Note « Méthodes et outils émergents d'évaluation des dangers ». (2016).
51. Ineris. Note de synthèse - L'utilisation des chemins de l'effet néfaste (AOP) dans le cadre de l'évaluation de la toxicité des substances chimiques. (2021).
52. Leist, M. *et al.* Adverse outcome pathways: opportunities, limitations and open questions. *Arch. Toxicol.* **91**, 3477–3505 (2017).
53. Council, N. R., Studies, D. on E. and L., Research, I. for L. A., Toxicology, B. on E. S. and Agents, C. on T. T. and A. of E. *Toxicity Testing in the 21st Century: A Vision and a Strategy*. (National Academies Press, 2007).
54. Zuang, V. *et al.* *Non-Animal Methods in Science and Regulation: EURL ECVAM Status Report 2021*. (Publications Office of the European Union, 2022). doi:10.2760/93290.
55. van der Zalm, A. J. *et al.* A framework for establishing scientific confidence in new approach methodologies. *Arch. Toxicol.* **96**, 2865–2879 (2022).
56. Westmoreland, C. *et al.* Use of New Approach Methodologies (NAMs) in regulatory decisions for chemical safety: Report from an EPAA Deep Dive Workshop. *Regul. Toxicol. Pharmacol.* **135**, 105261 (2022).
57. CNRS. La méthode idéale pour intégrer vos données multi-omiques existe-t-elle? | CNRS Biologie. <https://www.insb.cnrs.fr/fr/cnrsinfo/la-methode-ideale-pour-integrer-vos-donnees-multi-omiques-existe-t-elle> (2021).
58. Edmondson, R., Broglie, J. J., Adcock, A. F. & Yang, L. Three-Dimensional Cell Culture Systems and Their Applications in Drug Discovery and Cell-Based Biosensors. *Assay Drug Dev. Technol.* **12**, 207–218 (2014).
59. Lea, T. Caco-2 Cell Line. in *The Impact of Food Bioactives on Health: in vitro and ex vivo models* (eds. Verhoecx, K. et al.) (Springer, Cham (CH), 2015).
60. Ankley, G. T. *et al.* Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem.* **29**, 730–741 (2010).
61. Raies, A. B. & Bajic, V. B. In silico toxicology: computational methods for the prediction of chemical toxicity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **6**, 147–172 (2016).
62. Raies, A. B. & Bajic, V. B. In silico toxicology: computational methods for the prediction of chemical toxicity. *WIREs Comput. Mol. Sci.* **6**, 147–172 (2016).
63. Gégot, S. & Joubert, O. Le futur des expérimentations en toxicologie. *Environ. Risques Santé - ERS* 457–465 (2021).
64. Romano, J. D., Hao, Y. & Moore, J. H. Improving QSAR Modeling for Predictive Toxicology using Publicly Aggregated Semantic Graph Data and Graph Neural Networks. *Pac. Symp. Biocomput. Pac. Symp. Biocomput.* **27**, 187–198 (2022).
65. Chang, X. *et al.* IVIVE: Facilitating the Use of In Vitro Toxicity Data in Risk Assessment and Decision Making. *Toxics* **10**, 232 (2022).
66. Cronin, M. *et al.* The In Chemico-In Silico Interface: Challenges for Integrating Experimental and Computational Chemistry to Identify Toxicity. *Altern. Lab. Anim. ATLA* **37**, 513–21 (2009).
67. OECD. *Test No. 442C: In Chemico Skin Sensitisation: Assays Addressing the Adverse Outcome*

Pathway Key Event on Covalent Binding to Proteins. (Organisation for Economic Co-operation and Development, Paris, 2023).

68. OECD. Essai n° 442D: Sensibilisation cutanée in vitro: Méthode d'essai ARE-Nrf2 luciférase. <https://www.oecd.org/fr/securitechimique/essai-n-442d-sensibilisation-cutanee-in-vitro-9789264229778-fr.htm> (2022).

69. OECD. Essai n° 442E: Sensibilisation cutanée in vitro: Essai de sensibilisation cutanée in vitro portant sur l'événement clé relatif à l'activation des cellules dendritiques dans la voie toxicologique impliquée dans les effets indésirables pour la sensibilisation cutanée | READ online. https://read.oecd-ilibrary.org/environment/essai-n-442e-sensibilisation-cutanee-in-vitro_9789264276505-fr (2023).

70. Dusinska, M., Rundén-Pran, E., Carreira, S. C. & Saunders, M. Chapter 4 - Critical Evaluation of Toxicity Tests. in *Adverse Effects of Engineered Nanomaterials* (eds. Fadeel, B., Pietroiusti, A. & Shvedova, A. A.) 63–83 (Academic Press, Boston, 2012). doi:10.1016/B978-0-12-386940-1.00004-0.

71. Leushacke, M. & Barker, N. Ex vivo culture of the intestinal epithelium: strategies and applications. *Gut* **63**, 1345–1354 (2014).

72. OECD. *Test No. 437: Bovine Corneal Opacity and Permeability Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage*. (Organisation for Economic Co-operation and Development, Paris, 2023).

73. Claude, N., Goldfain-Blanc, F. & Guillouzo, A. La place des méthodes in silico, in vitro, in omic dans l'évaluation de la sécurité des médicaments. *médecine/sciences* **25**, 105–110 (2009).

74. OECD. Omics technologies in chemical testing - OECD. <https://www.oecd.org/chemicalsafety/testing/omics.htm>.

75. Gwinn, W. M. *et al.* Evaluation of 5-day In Vivo Rat Liver and Kidney With High-throughput Transcriptomics for Estimating Benchmark Doses of Apical Outcomes. *Toxicol. Sci.* **176**, 343–354 (2020).

76. Claude, N., Goldfain-Blanc, F. & Guillouzo, A. La place des méthodes in silico, in vitro, in omic dans l'évaluation de la sécurité des médicaments. *médecine/sciences* **25**, 105–110 (2009).

77. Hemmerich, J. & Ecker, G. F. In silico toxicology: From structure–activity relationships towards deep learning and adverse outcome pathways. *WIREs Comput. Mol. Sci.* **10**, e1475 (2020).

78. Archibald, K., Tsaïoun, K., Kenna, J. G. & Pound, P. Better science for safer medicines: the human imperative. *J. R. Soc. Med.* **111**, 433–438 (2018).

79. Yusuf, M. Insights into the in-silico research: Current scenario, advantages, limits, and future perspectives. *Life Silico* **1**, 13–25 (2023).

80. Lugli, G. A. *et al.* A microbiome reality check: limitations of in silico-based metagenomic approaches to study complex bacterial communities. *Environ. Microbiol. Rep.* **11**, 840–847 (2019).

81. Ineris. Les méthodes alternatives en expérimentation animale. (2013).

82. European Chemicals Agency (EU body or agency). *The Use of Alternatives to Testing on Animals for the REACH Regulation*. (Publications Office of the European Union, LU, 2020).

83. European Commission. The European Green Deal - European Commission. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en (2021).

84. Gourmelon, A. Toxicologie, évaluations réglementaires des dangers/risques sans tests sur les animaux - vers des outils communs à travers les différentes réglementations. (2023).

85. Pistollato, F. *et al.* Current EU regulatory requirements for the assessment of chemicals and cosmetic products: challenges and opportunities for introducing new approach methodologies. *Arch. Toxicol.* **95**, 1867–1897 (2021).

86. Van der Stel, W. *et al.* New approach methods (NAMs) supporting read-across: Two neurotoxicity AOP-based IATA case studies. *ALTEX* **38**, 615–635 (2021).

87. European Commission. Partnership for the Assessment of Risks from Chemicals | Parc. <https://www.eu-parc.eu/> (2024).
88. Westmoreland, C. *et al.* Use of New Approach Methodologies (NAMs) in regulatory decisions for chemical safety: Report from an EPAA Deep Dive Workshop. *Regul. Toxicol. Pharmacol.* **135**, 105261 (2022).
89. US EPA, O. New Approach Methods Work Plan, Reducing use of Animal testing. (2020).
90. Canada, E. and C. C. Bill S-5: Strengthening Environmental Protection for a Healthier Canada Act. <https://www.canada.ca/en/environment-climate-change/news/2023/06/bill-s-5-strengthening-environmental-protection-for-a-healthier-canada-act.html> (2023).
91. Understanding Animal Research. Cosmetic testing. *Understanding Animal Research* <https://www.understandinganimalresearch.org.uk/regulation/cosmetics>.
92. European Commission. The Commission roadmap for phasing out animal testing in chemical safety assessments - European Commission. https://single-market-economy.ec.europa.eu/events/commission-roadmap-phasing-out-animal-testing-chemical-safety-assessments-2023-12-11_en (2023).
93. European Chemicals Agency. *Report on the European Chemicals Agency's "New Approach Methodologies Workshop: Towards an Animal Free Regulatory System for Industrial Chemicals" (31 May - 1 June 2023, Helsinki, Finland)*. (Publications Office, LU, 2023).
94. European Chemicals Agency. *Key Areas of Regulatory Challenge*. (Publications Office, LU, 2023).
95. Pistollato, F. *et al.* Current EU regulatory requirements for the assessment of chemicals and cosmetic products: challenges and opportunities for introducing new approach methodologies. *Arch. Toxicol.* **95**, 1867–1897 (2021).
96. Kutsarova, S. *et al.* Automated read-across workflow for predicting acute oral toxicity: I. The decision scheme in the QSAR toolbox. *Regul. Toxicol. Pharmacol.* **125**, 105015 (2021).
97. Ball, N. *et al.* t4 report: Toward Good Read-Across Practice (GRAP) Guidance. *ALTEX* **33**, 149–166 (2016).
98. Cronin, M. T. D. An Introduction to Chemical Grouping, Categories and Read-Across to Predict Toxicity. (2013) doi:10.1039/9781849734400-00001.
99. OECD. Grouping of Chemicals: Chemical Categories and Read-Across. <https://web.archive.oecd.org/2014-04-16/66317-groupingofchemicalschemicalcategoriesandread-across.htm> (2014).
100. ECHA. Read-Across Assessment Framework (RAAF). (2017).
101. Schultz, T. W., Richarz, A.-N. & Cronin, M. T. D. Assessing uncertainty in read-across: Questions to evaluate toxicity predictions based on knowledge gained from case studies. *Comput. Toxicol.* **9**, 1–11 (2019).
102. Stepan, A. F. *et al.* Structural alert/reactive metabolite concept as applied in medicinal chemistry to mitigate the risk of idiosyncratic drug toxicity: a perspective based on the critical examination of trends in the top 200 drugs marketed in the United States. *Chem. Res. Toxicol.* **24**, 1345–1410 (2011).
103. Tropsha, A. & Golbraikh, A. Predictive QSAR modeling workflow, model applicability domains, and virtual screening. *Curr. Pharm. Des.* **13**, 3494–3504 (2007).
104. Mombelli, E. La modélisation QSAR. (2008).
105. ECHA. QSAR models - ECHA. <https://echa.europa.eu/support/registration/how-to-avoid-unnecessary-testing-on-animals/qsar-models>.
106. Ahlberg, E., Carlsson, L. & Boyer, S. Computational Derivation of Structural Alerts from Large Toxicology Data Sets. *J. Chem. Inf. Model.* **54**, 2945–2952 (2014).
107. Alves, V. *et al.* Alarms about structural alerts. *Green Chem. Int. J. Green Chem. Resour. GC* **18**, 4348–4360 (2016).

108. Mansouri, K. *et al.* CATMoS: Collaborative Acute Toxicity Modeling Suite. *Environ. Health Perspect.* **129**, 047013 (2021).
109. Saiakhov, R. D., Stefan, L. R. & Klopman, G. Multiple computer-automated structure evaluation model of the plasma protein binding affinity of diverse drugs. *Perspect. Drug Discov. Des.* **19**, 133–155 (2000).
110. European Commission, Joint Research Center (JRC). Joint Research Centre Data Catalogue - JRC QSAR Model Database. <https://data.jrc.ec.europa.eu/dataset/e4ef8d13-d743-4524-a6eb-80e18b58cba4> (2020).
111. Technical University of Denmark (DTU). Danish (Q)SAR Database. <http://qsar.food.dtu.dk/> (2004).
112. Dimitrov, S. D. *et al.* QSAR Toolbox – workflow and major functionalities. *SAR QSAR Environ. Res.* **27**, 203–219 (2016).
113. EUR-Lex. Council Regulation (EC) No. 440/2008 laying down test methods pursuant to Regulation (EC) No. 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). (2008).
114. Trevan, J. W. & Dale, H. H. The error of determination of toxicity. *Proc. R. Soc. Lond. Ser. B Contain. Pap. Biol. Character* **101**, 483–514 (1927).
115. Hamm, J. *et al.* Alternative approaches for identifying acute systemic toxicity: Moving from research to regulatory testing. *Toxicol. In Vitro* **41**, 245–259 (2017).
116. EUR-Lex. EUR-Lex - 02008R1272-20200501 - EN - EUR-Lex. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02008R1272-20200501> (2020).
117. ECHA. Guidance on Information Requirements and Chemical Safety Assessment, Chapter R.7a: Endpoint specific guidance Version 6.0. (2017).
118. ECHA. Guidance on the Application of the CLP Criteria - Guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures. (2017).
119. Dearden, J. C., Cronin, M. T. D. & Kaiser, K. L. E. How not to develop a quantitative structure–activity or structure–property relationship (QSAR/QSPR). *SAR QSAR Environ. Res.* **20**, 241–266 (2009).
120. Ekwall, B. *et al.* MEIC Evaluation of Acute Systemic Toxicity: Part VI. The Prediction of Human Toxicity by Rodent LD50 Values and Results from 61 In Vitro Methods. *Altern. Lab. Anim.* **26**, 617–658 (1998).
121. Zbinden, G. & Flury-Roversi, M. Significance of the LD50-test for the toxicological evaluation of chemical substances. *Arch. Toxicol.* **47**, 77–99 (1981).
122. EUR-Lex. Commission Recommendation of 18 June 2007 on Guidelines for the Accommodation and Care of Animals Used for Experimental and Other Scientific Purposes (Notified under Document Number C(2007) 2525) (Text with EEA Relevance). OJ L vol. 197 <http://data.europa.eu/eli/reco/2007/526/oj/eng> (2007).
123. Russell, W. M. S. & Burch, R. L. Prefatory Note. *Altern. Lab. Anim. ATLA* **37**, 267–268 (2009).
124. Prieto, P. *et al.* Investigating cell type specific mechanisms contributing to acute oral toxicity. *ALTEX - Altern. Anim. Exp.* **36**, 39–64 (2019).
125. QSAR Toolbox 4.6. Databases. *QSAR Toolbox* <https://qsartoolbox.org/resources/databases/>.
126. OCDE. A propos de l'Organisation de Coopération et de Développement Économiques (OCDE) - OCDE. <https://www.oecd.org/fr/apropos/aproposdelorganisationdecooperationetdedeveloppementeconomiquesocde.htm> (2012).
127. OECD. The OECD QSAR Toolbox - OECD. <https://www.oecd.org/chemicalsafety/risk-assessment/oecd-qsar-toolbox.htm>.
128. OECD. Application manual of OECD QSAR Toolbox v.4. (2017).
129. Kuseva, C. *et al.* Category consistency in the OECD QSAR Toolbox: Assessment and reporting tool to justify read-across. *Comput. Toxicol.* **11**, 65–71 (2019).

130. Nikolov, N. G. & Wedebye, E. B. Recent developments of the Danish (Q)SAR Database: Updates, statistics and link to the OECD QSAR Toolbox: QSAR2018. (2018).
131. National Library of Medicine. ChemIDplus and the Drug Information Portal Content Available From PubChem Only Starting December 2022. <https://www.nlm.nih.gov/pubs/techbull/tb.html> (2022).
132. Nedelcheva, D. *et al.* In silico mechanistically-based profiling module for acute oral toxicity. *Comput. Toxicol.* **12**, 100109 (2019).
133. Myatt, G. J. *et al.* In silico toxicology protocols. *Regul. Toxicol. Pharmacol.* **96**, 1–17 (2018).
134. Ankley, G. T. *et al.* Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem.* **29**, 730–741 (2010).
135. Hadrup, N. *et al.* Asthma-inducing potential of 28 substances in spray cleaning products—Assessed by quantitative structure activity relationship (QSAR) testing and literature review. *J. Appl. Toxicol.* **42**, 130–153 (2022).
136. Wedebye, E. B. User Manual for the Danish (Q)SAR Database. (2018).
137. Vayer, P., Arrault, A., Lesur, B., Bertrand, M. & Walther, B. Apports de la chémo-informatique dans la recherche et l'optimisation des molécules d'intérêt thérapeutique. *médecine/sciences* **25**, 871–877 (2009).
138. Silberman, J. & Taylor, A. Carbamate Toxicity. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2023).
139. Vale, A. & Lotti, M. Chapter 10 - Organophosphorus and carbamate insecticide poisoning. in *Handbook of Clinical Neurology* (eds. Lotti, M. & Bleecker, M. L.) vol. 131 149–168 (Elsevier, 2015).
140. Abd-Ella, A. ETUDE DU MODE D'ACTION NEUROTOXIQUE D'UN REPULSIF, LE DEET UTILISE SEUL ET EN ASSOCIATION AVEC UN INSECTICIDE SUR L'ACÉTYLCHOLINESTERASE DES DUM NEURONES D'UN INSECTE LA Spécialité : Sciences Agronomique. (2011).
141. Bauer, F., Thomas, P., Fouchard, S. & Serge, N. High-accuracy prediction of Mechanisms of Action using structural alerts. *Comput. Toxicol.* **7**, (2018).
142. Tadeo, J. L. & Sánchez-Brunete, C. HERBICIDES | Gas Chromatography☆. in *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering* (Elsevier, 2013). doi:10.1016/B978-0-12-409547-2.04775-2.
143. Munro, I. C., Ford, R. A., Kennepohl, E. & Sprenger, J. G. Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. *Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc.* **34**, 829–867 (1996).
144. Cramer, G. M., Ford, R. A. & Hall, R. L. Estimation of toxic hazard—a decision tree approach. *Food Cosmet. Toxicol.* **16**, 255–276 (1978).
145. Gulea, M. Progress in the Chemistry of Phosphorothioates. *Adv. Org. Synth.* **12**, 117–150 (2018).
146. Vale, A. & Lotti, M. Chapter 10 - Organophosphorus and carbamate insecticide poisoning. in *Handbook of Clinical Neurology* (eds. Lotti, M. & Bleecker, M. L.) vol. 131 149–168 (Elsevier, 2015).
147. Zhou, Y. *et al.* Aryl-, halogenated- and alkyl-organophosphate esters induced oxidative stress, endoplasmic reticulum stress and NLRP3 inflammasome activation in HepG2 cells. *Environ. Pollut.* **316**, 120559 (2023).
148. INRS. Organophosphorés (-) / Cholinestérasés intraérythrocytaires - Biotox - INRS. https://www.inrs.fr/publications/bdd/biotox/dosage.html?refINRS=Dosage_122 (2003).
149. Blum, A. *et al.* Organophosphate Ester Flame Retardants: Are They a Regrettable Substitution for Polybrominated Diphenyl Ethers? *Environ. Sci. Technol. Lett.* **6**, 638–649 (2019).
150. Gbadamosi, M. R., Abdallah, M. A.-E. & Harrad, S. A critical review of human exposure to organophosphate esters with a focus on dietary intake. *Sci. Total Environ.* **771**, 144752 (2021).

151. Bauer, F. Une meilleure caractérisation des mécanismes d'action toxique à partir de la structure moléculaire. (Université Haute-Alsace, 2017).
152. Colovic, M. B., Krstic, D. Z., Lazarevic-Pasti, T. D., Bondzic, A. M. & Vasic, V. M. Acetylcholinesterase Inhibitors: Pharmacology and Toxicology. *Curr. Neuropharmacol.* **11**, 315–335 (2013).
153. Myatt, G. J. *et al.* Increasing the acceptance of in silico toxicology through development of protocols and position papers. *Comput. Toxicol.* **21**, 100209 (2022).
154. Karmaus, A. L. *et al.* Evaluation of Variability Across Rat Acute Oral Systemic Toxicity Studies. *Toxicol. Sci.* **188**, 34–47 (2022).
155. ECHA. Guidance on the Application of the CLP Criteria - Guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures. (2017).
156. Aspis. Aspis - Project cluster for Implementation of novel Strategies. <https://aspis-cluster.eu/> (2024).
157. US EPA, O. Toxicology Testing in the 21st Century (Tox21). <https://www.epa.gov/chemical-research/toxicology-testing-21st-century-tox21> (2015).
158. Van, T. Artificial Intelligence in Drug Toxicity Prediction: Recent Advances, Challenges, and Future Perspectives. *J. Chem. Inf. Model.* **63**, (2023).
159. Wang, M. W. H., Goodman, J. M. & Allen, T. E. H. Machine Learning in Predictive Toxicology: Recent Applications and Future Directions for Classification Models. *Chem. Res. Toxicol.* **34**, 217–239 (2021).
160. Hartung, T. ToxAlcology - The evolving role of artificial intelligence in advancing toxicology and modernizing regulatory science. *ALTEX - Altern. Anim. Exp.* **40**, 559–570 (2023).
161. Kavlock, R. *et al.* Update on EPA's ToxCast Program: Providing High Throughput Decision Support Tools for Chemical Risk Management. *Chem. Res. Toxicol.* **25**, 1287–1302 (2012).
162. Gilmour, N. *et al.* Next generation risk assessment for skin allergy: Decision making using new approach methodologies. *Regul. Toxicol. Pharmacol.* **131**, 105159 (2022).
163. Kleinstreuer, N. C. *et al.* Predictive models for acute oral systemic toxicity: A workshop to bridge the gap from research to regulation. *Comput. Toxicol.* **8**, 21–24 (2018).
164. Esteva, A. *et al.* A guide to deep learning in healthcare. *Nat. Med.* **25**, 24–29 (2019).
165. Hartung, T. Artificial intelligence as the new frontier in chemical risk assessment. *Front. Artif. Intell.* **6**, 1269932 (2023).
166. Elton, D. C., Boukouvalas, Z., Fuge, M. D. & Chung, P. W. Deep learning for molecular design—a review of the state of the art. *Mol. Syst. Des. Eng.* **4**, 828–849 (2019).
167. Cavasotto, C. N. & Scardino, V. Machine Learning Toxicity Prediction: Latest Advances by Toxicity End Point. *ACS Omega* **7**, 47536–47546 (2022).
168. Lin, Z. & Chou, W.-C. Machine Learning and Artificial Intelligence in Toxicological Sciences. *Toxicol. Sci.* **189**, 7–19 (2022).
169. Shah, A. & Jain, M. Chapter 9 - Limitations and future challenges of computer-aided drug design methods. in *Computer Aided Drug Design (CADD): From Ligand-Based Methods to Structure-Based Approaches* (eds Rudrapal, M. & Egbuna, C.) 283–297 (Elsevier, 2022). doi:10.1016/B978-0-323-90608-1.00006-X.
170. Tsou, L. K. *et al.* Comparative study between deep learning and QSAR classifications for TNBC inhibitors and novel GPCR agonist discovery. *Sci. Rep.* **10**, 16771 (2020).
171. Staszak, M. *et al.* Machine learning in drug design: Use of artificial intelligence to explore the chemical structure–biological activity relationship. *WIREs Comput. Mol. Sci.* **12**, e1568 (2022).
172. Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E. & Svetnik, V. Deep Neural Nets as a Method for Quantitative Structure–Activity Relationships. *J. Chem. Inf. Model.* **55**, 263–274 (2015).

173. Zhao, Y., Xia, Y., Yu, Y. & Liang, G. QSAR in natural non-peptidic food-related compounds: Current status and future perspective. *Trends Food Sci. Technol.* **140**, 104165 (2023).
174. Mayr, A., Klambauer, G., Unterthiner, T. & Hochreiter, S. DeepTox: Toxicity Prediction using Deep Learning. *Front. Environ. Sci.* **3**, (2016).
175. Garrido, A., Lepailleur, A., Mignani, S. M., Dallemagne, P. & Rochais, C. hERG toxicity assessment: Useful guidelines for drug design. *Eur. J. Med. Chem.* **195**, 112290 (2020).
176. Sanguinetti, M. C., Jiang, C., Curran, M. E. & Keating, M. T. A mechanistic link between an inherited and an acquired cardiac arrhythmia: HERG encodes the IKr potassium channel. *Cell* **81**, 299–307 (1995).
177. Venkatraman, V. FP-ADMET: a compendium of fingerprint-based ADMET prediction models. *Cheminformatics* **13**, 75 (2021).
178. Hinselmann, G., Rosenbaum, L., Jahn, A., Fechner, N. & Zell, A. jCompoundMapper: An open source Java library and command-line tool for chemical fingerprints. *J. Cheminformatics* **3**, 3 (2011).
179. Cai, C. *et al.* Deep Learning-Based Prediction of Drug-Induced Cardiotoxicity. *J. Chem. Inf. Model.* **59**, 1073–1084 (2019).
180. Zhang, X., Mao, J., Wei, M., Qi, Y. & Zhang, J. Z. H. HergSPred: Accurate Classification of hERG Blockers/Nonblockers with Machine-Learning Models. *J. Chem. Inf. Model.* **62**, 1830–1839 (2022).
181. Kumachev, A. & Wu, P. E. Atteinte hépatique médicamenteuse. *CMAJ Can. Med. Assoc. J.* **193**, E746–E747 (2021).
182. Babai, S., Auclert, L. & Le-Louët, H. Safety data and withdrawal of hepatotoxic drugs. *Therapies* **76**, 715–723 (2021).
183. Vall, A. *et al.* The Promise of AI for DILI Prediction. *Front. Artif. Intell.* **4**, (2021).
184. O'Brien, P. J. *et al.* High concordance of drug-induced human hepatotoxicity with in vitro cytotoxicity measured in a novel cell-based model using high content screening. *Arch. Toxicol.* **80**, 580–604 (2006).
185. Chen, M. *et al.* DILIRank: the largest reference drug list ranked by the risk for developing drug-induced liver injury in humans. *Drug Discov. Today* **21**, 648–653 (2016).
186. Kotsampasakou, E., Montanari, F. & Ecker, G. F. Predicting drug-induced liver injury: The importance of data curation. *Toxicology* **389**, 139–145 (2017).
187. Richard, A. M. *et al.* The Tox21 10K Compound Library: Collaborative Chemistry Advancing Toxicology. *Chem. Res. Toxicol.* **34**, 189–216 (2021).
188. Béquignon, O. J. M., Pawar, G., van de Water, B., Cronin, M. T. D. & van Westen, G. J. P. Computational Approaches for Drug-Induced Liver Injury (DILI) Prediction: State of the Art and Challenges. in *Systems Medicine* (ed. Wolkenhauer, O.) 308–329 (Academic Press, Oxford, 2021). doi:10.1016/B978-0-12-801238-3.11535-1.
189. Pizzo, F., Lombardo, A., Manganaro, A. & Benfenati, E. A New Structure-Activity Relationship (SAR) Model for Predicting Drug-Induced Liver Injury, Based on Statistical and Expert-Based Structural Alerts. *Front. Pharmacol.* **7**, (2016).
190. King, A. M. & Aaron, C. K. Organophosphate and carbamate poisoning. *Emerg. Med. Clin. North Am.* **33**, 133–151 (2015).
191. Trang, A. & Khandhar, P. B. Physiology, Acetylcholinesterase. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2023).
192. Wilson, R. S. *et al.* The natural history of cognitive decline in Alzheimer's disease. *Psychol. Aging* **27**, 1008–1017 (2012).
193. Sandhu, H., Kumar, R. N. & Garg, P. Machine learning-based modeling to predict inhibitors of acetylcholinesterase. *Mol. Divers.* **26**, 331–340 (2022).
194. Whitehouse, P. J. Cholinergic therapy in dementia. *Acta Neurol. Scand.* **88**, 42–45 (1993).
195. Huang, J. Maladie d'Alzheimer - Troubles neurologiques. *Édition professionnelle du Manuel MSD*
<https://www.msmanuals.com/fr/professional/troubles-neurologiques/syndrome-confusionnel-et->

d%C3%A9mence/maladie-alzheimer?query=rivastigmine (2023).

196. Sapundzhi, F., Dzimbova, T., Pencheva, N. & Milanov, P. Comparative evaluation of four scoring functions with three models of delta opioid receptor using molecular docking. (2016).

197. Khan, M. I. *et al.* Discovery of novel acetylcholinesterase inhibitors through integration of machine learning with genetic algorithm based in silico screening approaches. *Front. Neurosci.* **16**, (2023).

198. WHO. Pharmacovigilance. <https://www.who.int/teams/regulation-prequalification/regulation-and-safety/pharmacovigilance>.

199. Géniaux, H. & Laroche, M.-L. La pharmacovigilance, fonctionnement et missions. *Actual. Pharm.* **56**, 20–23 (2017).

200. Kompa, B. *et al.* Artificial Intelligence Based on Machine Learning in Pharmacovigilance: A Scoping Review. *Drug Saf.* **45**, 477–491 (2022).

201. Bate, A. & Hobbiger, S. F. Artificial Intelligence, Real-World Automation and the Safety of Medicines. *Drug Saf.* **44**, 125–132 (2021).

202. ANSM. Actualité - Lancement de la nouvelle application nationale de pharmacovigilance. ANSM <https://ansm.sante.fr/actualites/lancement-de-la-nouvelle-application-nationale-de-pharmacovigilance> (2021).

203. Jensen, P. B., Jensen, L. J. & Brunak, S. Mining electronic health records: towards better research applications and clinical care. *Nat. Rev. Genet.* **13**, 395–405 (2012).

204. Yadav, P., Steinbach, M., Kumar, V. & Simon, G. Mining Electronic Health Records (EHRs): A Survey. *ACM Comput. Surv.* **50**, 85:1-85:40 (2018).

205. Ball, R., Robb, M., Anderson, S. & Dal Pan, G. The FDA's sentinel initiative—A comprehensive approach to medical product surveillance. *Clin. Pharmacol. Ther.* **99**, 265–268 (2016).

206. Desai, R. J. *et al.* Broadening the reach of the FDA Sentinel system: A roadmap for integrating electronic health record data in a causal analysis framework. *Npj Digit. Med.* **4**, 1–6 (2021).

207. Maiti, S. Extracting Medical Information From Clinical Text With NLP. *Analytics Vidhya* <https://www.analyticsvidhya.com/blog/2023/02/extracting-medical-information-from-clinical-text-with-nlp/> (2023).

208. Davis, S. E. *et al.* Use of Electronic Health Record Data for Drug Safety Signal Identification: A Scoping Review. *Drug Saf.* **46**, 725–742 (2023).

209. Hoffmann, S. & Hartung, T. Toward an evidence-based toxicology. *Hum. Exp. Toxicol.* **25**, 497–513 (2006).

210. Pilipiec, P., Liwicki, M. & Bota, A. Using Machine Learning for Pharmacovigilance: A Systematic Review. *Pharmaceutics* **14**, 266 (2022).

211. Glaser, M. *et al.* Automating individual case safety report identification within scientific literature using natural language processing. in vol. 30 118–118 (WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2021).

212. Raja, K., Patrick, M., Elder, J. T. & Tsoi, L. C. Machine learning workflow to enhance predictions of Adverse Drug Reactions (ADRs) through drug-gene interactions: application to drugs for cutaneous diseases. *Sci. Rep.* **7**, 3690 (2017).

213. Wang, X., Xu, X., Tong, W., Roberts, R. & Liu, Z. InferBERT: A Transformer-Based Causal Inference Framework for Enhancing Pharmacovigilance. *Front. Artif. Intell.* **4**, (2021).

214. Inserm. Intelligence artificielle et santé · Inserm, La science pour la santé. *Inserm* <https://www.inserm.fr/dossier/intelligence-artificielle-et-sante/> (2022).

215. Martineau, J. T. & Godin, F. R. Tour d'horizon des enjeux éthiques liés à l'IA en santé. *Éthique Publique Rev. Int. D'éthique Sociétale Gouv.* (2023) doi:10.4000/ethiquepublique.7978.

216. Vayena, E., Dzenowagis, J., Brownstein, J. S. & Sheikh, A. Policy implications of big data in the health sector. *Bull. World Health Organ.* **96**, 66–68 (2018).

217. Hand, D. J. Aspects of Data Ethics in a Changing World: Where Are We Now? *Big Data* **6**, 176–190 (2018).

218. de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M. & Blondel, V. D. Unique in the Crowd: The privacy bounds of human mobility. *Sci. Rep.* **3**, 1376 (2013).
219. WHO. Ethics and governance of artificial intelligence for health. <https://www.who.int/publications-detail-redirect/9789240029200> (2021).
220. DGOS. La cybersécurité: un enjeu majeur pour les établissements de santé. *Ministère de la Santé et de la Prévention* <https://sante.gouv.fr/systeme-de-sante/e-sante/sih/dossier-cybersecurite/article/la-cybersecurite-un-enjeu-majeur-pour-les-etablissements-de-sante> (2024).
221. Reddy, S., Allan, S., Coghlan, S. & Cooper, P. A governance model for the application of AI in health care. *J. Am. Med. Inform. Assoc.* **27**, 491–497 (2020).
222. Someh, I., Davern, M., Breidbach, C. & Shanks, G. Ethical Issues in Big Data Analytics: A Stakeholder Perspective. *Commun. Assoc. Inf. Syst.* **44**, (2019).
223. CNIL. Explicabilité (IA). <https://www.cnil.fr/fr/definition/explicabilite-ia>.
224. Birhane, A. Algorithmic injustice: a relational ethics approach. *Patterns* **2**, (2021).
225. Schneiderman, L. J. *et al.* Dissatisfaction with Ethics Consultations: The Anna Karenina Principle. *Camb. Q. Healthc. Ethics* **15**, 101–106 (2006).
226. Bender, A. 'AI' in Toxicology (In Silico Toxicology) – The Pieces Don't Yet Fit Together – DrugDiscovery.NET – AI in Drug Discovery. <http://www.drugdiscovery.net/2019/08/31/ai-in-toxicology-in-silico-toxicology-the-pieces-dont-yet-fit-together/> (2019).
227. Phillips, S. P., Spithoff, S. & Simpson, A. L'intelligence artificielle et les algorithmes prédictifs en médecine. *Can. Fam. Physician* **68**, e230–e233 (2022).
228. Kleinstreuer, N. & Hartung, T. Artificial intelligence (AI)—it's the end of the tox as we know it (and I feel fine)*. *Arch. Toxicol.* (2024) doi:10.1007/s00204-023-03666-2.
229. Kwon, Y., Won, J.-H., Kim, B. J. & Paik, M. C. Uncertainty quantification using Bayesian neural networks in classification: Application to biomedical image segmentation. *Comput. Stat. Data Anal.* **142**, 106816 (2020).
230. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
231. European Commission. Ontology-driven and artificial intelligence-based repeated dose toxicity testing of chemicals for next generation risk assessment | ONTOX Project | Fact Sheet | H2020. *CORDIS* | *European Commission* <https://cordis.europa.eu/project/id/963845> (2020).
232. ICCVAM. Validation, Qualification, and Regulatory Acceptance of New Approach Methodologies: draft August 2023. (2023).
233. OECD. *(Q)SAR Assessment Framework*. (2023).
234. Schaafsma, G., Kroese, E. D., Tielemans, E. L. J. P., Van de Sandt, J. J. M. & Van Leeuwen, C. J. REACH, non-testing approaches and the urgent need for a change in mind set. *Regul. Toxicol. Pharmacol.* **53**, 70–80 (2009).

LA PERTINENCE DES DONNÉES FOURNIES PAR LES MÉTHODES ALTERNATIVES À L'EXPERIMENTATION ANIMALE DANS L'ÉVALUATION DU DANGER EN TOXICOLOGIE : ÉTAT DES LIEUX, ÉTUDES DE CAS ET PERSPECTIVES.

RÉSUMÉ

Les méthodes traditionnelles d'évaluation du danger en toxicologie, effectuées sur les animaux, se voient depuis plusieurs décennies remises en question pour des raisons éthiques, pratiques et économiques, favorisant l'émergence de méthodes alternatives. Ces dernières, également appelées « Nouvelles Approches Méthodologiques » (NAM) en toxicologie, incluent majoritairement des modélisations mathématiques (in silico) et des tests sur les cellules (in vitro). Le présent document a pour objectif d'en évaluer la justesse, en prenant pour exemple une étude de cas sur un critère d'évaluation particulier, la toxicité orale aiguë. La méthode retenue consiste à comparer les valeurs expérimentales obtenues sur les animaux pour différents groupes de substances avec les données obtenues à l'aide de deux approches NAM : les read-across et la méthode QSAR (« Quantitative Structure-Activity Relationship »). Cette comparaison vise à déterminer si les approches NAM influencent l'évaluation des risques, et donc la classification selon les critères « CLP » (« Classification, Labelling and Packaging ») propre à la toxicité aiguë, en vue de l'élaboration de directives d'essai par les autorités européennes. Cette étude permet de montrer la pertinence croissante des NAM, qui devrait s'intensifier avec les progrès effectués dans le domaine de l'intelligence artificielle (IA). Cette dernière offre la possibilité d'appréhender la complexité des ensembles de données liées à la toxicité, et peut être appliquée à de nombreux aspects de la toxicologie prédictive, contribuant ainsi à la réduction de l'expérimentation animale. Malgré des avantages prometteurs, des défis persistent quant à l'adoption de ces nouvelles méthodes par les autorités réglementaires. Elles devront faire l'objet de futures collaborations scientifiques et répondre à des exigences éthiques et réglementaires. Ces nouvelles méthodes alternatives, appliquées à l'évaluation du danger en toxicologie, pourraient significativement transformer la réglementation entourant les substances chimiques.

EVALUATION OF DATA RELEVANCE FROM ALTERNATIVES TO ANIMAL TESTING IN TOXICOLOGICAL HAZARD ASSESSMENT: CURRENT STATUS, CASE STUDIES, AND FUTURE PERSPECTIVES.

ABSTRACT

Traditional methods of toxicological hazard assessment, which rely on animals testing, have faced scrutiny for several decades due to ethical, practical, and economic reasons, promoting the emergence of alternative methods. These alternative methods, also known as "New Approach Methodologies" (NAM) in toxicology, primarily encompass mathematical modelling (in silico) and cell testing (in vitro). The aim of this document is to evaluate their accuracy, illustrated through case studies focusing on acute oral toxicity. The methodology involves comparing experimental data obtained from animal studies across different substance groups, with results derived from two NAM approaches: read-across and QSAR ("Quantitative Structure-Activity Relationship"). The objective of this comparison is to ascertain whether NAM approaches influence risk assessment, particularly in determining the "CLP" classification, specific to acute toxicity, for the development of test guidelines by European authorities. This study underscores the growing importance of NAM, which is expected to gain momentum alongside advancements in artificial intelligence (AI). AI presents opportunities for comprehending the complexities of toxicity-related datasets and can be applied across various facets of predictive toxicology, thereby contributing to the reduction of animal testing. Despite promising advantages, challenges persist in the adoption of these new methods by regulatory bodies, calling for future scientific collaborations and adherence to ethical and regulatory standards. These alternative approaches to hazard assessment in toxicology could significantly alter the regulation of chemical substances.

Valentine De Ketelaere