Joint Research Programme
BTO 2023.060 | Augustus 2023

Zijn persistente mobiele stoffen minder giftig?

Koppeling van fysisch-chemische stofeigenschappen aan gemeten toxiciteit

Report

Zijn persistente mobiele stoffen minder giftig?

Koppeling van fysisch-chemische stofeigenschappen aan gemeten toxiciteit

BTO 2023.060 | Augustus 2023

This research is part of the Joint Research Programme of KWR, the water utilities and Vewin.

Project number
402045-156

Project manager
Patrick Bauerlein

Client
BTO - Bedrijfsonderzoek

Author(s)
Milo de Baat, Renske Hoondert, Tessa Pronk, Thomas ter Laak

Quality Assurance
Milou Dingemans

Sent to
This report is distributed to BTO-participants and is public.

Keywords

PMOC, toxicology, exposure, toxic mechanisms, physicochemical properties of organic microcontaminants, computational toxicology

Year of publishing
2023

More information
dr. Renske Hoondert of dr. Milo de Baat
T 030-6069709 of 030-6069756
E Renske.Hoonder@kwrwater.nl
Milo.de.Baar@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

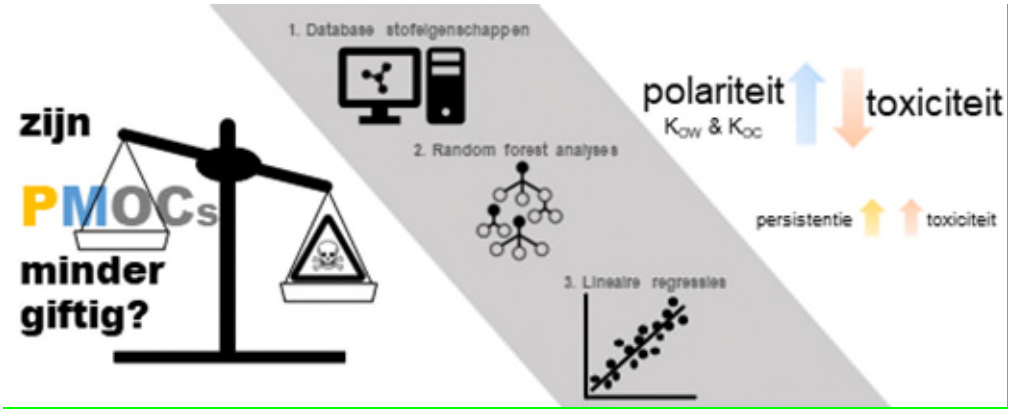
T +31 (0)30 60 69 511
E info@kwrwater.nl
of I www.kwrwater.nl

Managementsamenvatting

Zijn persistente mobiele organische stoffen minder giftig?

Auteurs: Milo de Baat, Renske Hoondert, Tessa Pronk, Thomas ter Laak

Zijn persistente mobiele organische stoffen minder giftig dan stoffen met andere eigenschappen? De persistentie van stoffen lijkt geen significante relatie met toxiciteit te hebben, maar meer mobiele stoffen blijken significant minder giftig dan minder mobiele stoffen. Dit is echter geen reden om voor deze stoffen minder aandacht te hebben, omdat de mobiele eigenschappen van deze stoffen uiteindelijk wel kunnen leiden tot hogere blootstelling en niet door een eenvoudige zuivering te verwijderen zijn.



Visualisatie van onderzoeksproject

Belang: PMOC zijn overal in het watersysteem
Persistente Mobiele Organische Chemicaliën (PMOC) lossen goed op in water en worden niet snel afgebroken. Daardoor verspreiden ze zich makkelijk door het watermilieu, en zijn ze moeilijk uit water te verwijderen. Mens en milieu kunnen dus via water in aanraking komen met PMOC. Dit maakt het essentieel de toxiciteit van deze stoffen beter in beeld te krijgen.

Aanpak: de relatie tussen PMOC en toxiciteit
De relatie tussen de stofeigenschappen die stoffen persistent en mobiel maken en hun toxiciteit is met statistische technieken onderzocht. Met *random forest* analyses is bestudeerd welke stof-eigenschappen correleren met gemeten effectconcentraties in celtesten (een maat voor toxiciteit). Vervolgens is met *lineaire regressiemodellen* in de ToxCast dataset met 111 toxiciteitstesten en 2845 stoffen de relatie tussen deze stofeigenschappen en effectconcentraties geanalyseerd.

Resultaten: mobiele stoffen zijn gemiddeld minder giftig, persistentie zegt daarover niets
De *random forest analyse* toonde een verband tussen toxiciteit en diverse stofkenmerken aan. De daaropvolgende *regressieanalyses* lieten een significante correlatie zien tussen toxiciteit en stofeigenschappen die de mobiliteit bepalen, zowel voor de gehele dataset als voor subsets van specifieke toxicologische eindpunten of stofgroepen. Daarbij bleken mobiele stoffen gemiddeld minder giftig. Voor de biodegradatie halfwaardetijd (een maat voor de persistentie van een stof) werden zulke verbanden niet gevonden.

Conclusie: mobiele stoffen blijken minder giftig, maar zijn daarom nog niet minder relevant
Mobiele stoffen zijn dus (gemiddeld) minder giftig, terwijl de persistentie van stoffen geen significante

relatie met toxiciteit laat zien. Dit betekent dat toxicologisch onderzoek naar persistente stoffen niet prioritair is aan onderzoek naar niet-persistente stoffen, mits natuurlijk de blootstelling niet significant verschilt. Dit wil echter niet zeggen dat de aanwezigheid van mobiele stoffen geen probleem is. Stoffen die persistent zijn kunnen doordringen in de waterketen en ophopen in het (water)milieu en een voortdurende blootstelling via (drink)water veroorzaken. Bovendien zijn de in dit onderzoek gebruikte toxiciteitstesten niet direct te vertalen naar effecten op het niveau van het intacte organisme (in dit geval de mens). De vraag rijst of *in vitro* assays wel de juiste testen zijn om daadwerkelijke effecten op organismeniveau te bepalen. Mogelijk leiden *in vitro* assays tot te hoge of juiste te lage schattingen van het effect van stoffen. Meer onderzoek is dan ook nodig om te bepalen of toxiciteitsgegevens op basis van *in vitro* assays wel geschikt zijn in de prioritering van (PMOC)-stoffen.

Toepassing: PMOC relevant voor onderzoek én beleid

Uit het huidige onderzoek is gebleken dat met name mobiliteit een (omgekeerd evenredige) significante relatie vertoont met toxiciteit (Een hogere mobiliteit van de stof lijkt een lagere toxiciteit te veroorzaken). Desalniettemin kan een hogere persistentie van de stof zorgen een uiteindelijk hogere blootstelling en wellicht een hoger risico voor de volksgezondheid. De combinatie van persistentie en mobiliteit zorgt immers voor voortdurende blootstelling via de waterketen en slechts een deel van de toxiciteit verklaard worden met eigenschappen die de mobiliteit beïnvloeden. Hoewel uit dit onderzoek dan ook niet blijkt dat PMOCs prioritaire stoffen zijn, blijft het wel van belang om de stoffen voldoende te blijven monitoren, gezien hun persistentie. Dit vraagt om verdiepende studies naar de mechanismen om de relatie tussen specifiek stofeigenschappen en toxiciteit beter te begrijpen en te kunnen voorspellen.

Rapport

Dit onderzoek is beschreven in het rapport *Zijn persistente mobiele stoffen minder giftig? Koppeling van fysisch-chemische stofeigenschappen aan gemeten toxiciteit* (BTO 2023.060). In een vervolgonderzoek '*Deeper understanding of PMOC toxicity*' wordt de relatie tussen eigenschappen die de stoffen persistent en mobiel maken in verband gebracht met diverse toxicologische werkingsmechanismen.

Contents

1 Zijn persistente mobiele stoffen minder giftig?
Nederlandse samenvatting

1.1 Introductie

Historisch heeft de milieuchemie zich de afgelopen halve eeuw vooral gericht op hydrofobe organische stoffen en anorganische stoffen. Dit heeft deels te maken met de beschikbaarheid van analytische technieken zoals atomic absorption spectrometry (AAS) voor (zware) metalen en anorganische verbindingen en gaschromatografische technieken om vluchtige organische stoffen van (complexe) milieumatrices te scheiden (Reemtsma et al., 2016). Onderzoek heeft duidelijk laten zien dat organische hydrofobe stoffen een probleem kunnen vormen voor het ecosysteem, omdat ze de neiging hebben te accumuleren in sedimenten en bodems en daar lang aanwezig kunnen blijven, sterk accumuleren in organismen en daar toxische effecten kunnen veroorzaken. Dit heeft geleid tot regulering van stoffen met hoge persistentie (P), bioaccumulatie en sorptie aan sediment en bodem (B) en toxiciteit (T). Dit worden *PBT-stoffen* genoemd (European Commission, 2022a). Door deze regulering moeten alle industriële stoffen die worden ontwikkeld voldoen aan bepaalde criteria voordat ze voor bepaalde doeleinden op de Europese markt worden toegelaten. De eigenschappen van het brede pallet van stoffen dat wordt geproduceerd door de industrie is daarmee waarschijnlijk verschoven naar die van meer polaire stoffen, omdat deze stoffen minder problematisch werden geacht. Deze stoffen werden tot voor kort echter ook minder gemeten in diverse milieucompartimenten (Reemtsma et al., 2016), of getoetst in toxiciteitstudies. De polaire eigenschappen van deze stoffen maken dat ze doorgaans goed oplossen in water, minder sterk sorberen aan geosorbentia als bodem en sediment of sorbentia toegepast in waterbehandeling zoals actieve kool. Dit betekent ook dat ze zich makkelijker bewegen door de waterketen (mobieler zijn). In navolging zijn naast de *PBT-stoffen* dan ook de *PMT-stoffen* benoemd, waarbij M voor mobiel staat (Neumann & Schliebner, 2019). Als deze stoffen naast hun polaire eigenschappen ook nog persistent zijn, wat wil zeggen dat ze niet snel worden omgezet in andere stoffen of gemineraliseerd, dan zullen zowel natuurlijke processen in de (stedelijke) waterketen als drinkwaterzuiveringstechnieken niet of minder goed in staat zijn deze stoffen uit het water te verwijderen. Dit maakt dat mobiele persistente stoffen na emissie in het milieu grote kans hebben om in (bronnen van) drinkwater terecht te komen waardoor de mens via drinkwater kan worden blootgesteld (Sjerps et al. (2016).

De grote kans op blootstelling maakt het relevant om ook te onderzoeken of de eigenschappen van organische stoffen, die maken dat stoffen mobiel en persistent zijn in de waterketen en lastig te verwijderen zijn uit water met diverse (drink)waterzuiveringstechnieken, ook een relatie hebben met de intrinsieke giftigheid (toxiciteit) van deze stoffen. De onderzochte hypothese is dat de meeste toxicologische werkingsmechanismen van stoffen een gevolg zijn van een interactie (binding) van stoffen met biologische structuren (zoals celmembranen of eiwitstructuren). Bovendien hebben zeer mobiele stoffen mogelijk minder de neiging om op te hopen in organismen zoals mensen en daar dergelijke bindingen aan te gaan, en daarmee minder toxisch zouden zijn. Hier wordt onderzocht of dezelfde eigenschappen van deze stoffen die maken dat ze ver doordringen in de stedelijke waterketen en lastig te verwijderen zijn (en dus zorgen voor hoge blootstelling deze stoffen) ze gelijktijdig mogelijk minder giftig maken. Dit betekent dat hoewel de humane blootstelling aan deze mobiele stoffen voortdurend is, de intrinsieke toxiciteit (het effect van deze stof bij beschikbare interne concentratie) lager is. Dat zou betekenen dat gezondheidsrisico's, de resultante van blootstelling en effect, toch beperkt zou kunnen blijven.

1.2 Onderzoeksvraag

Met deze studie onderzoeken of en hoe PMOC-stofeigenschappen (persistente, mobiele en organische stoffen) die invloed hebben op de blootstelling aan de stof via drinkwater ook een relatie hebben met hun intrinsieke toxiciteit.

1.3 Experimentele opzet

Om dit te onderzoeken is een grootschalige data-analyse gedaan, waarbij de toxiciteit van stoffen, gemeten in laboratoriumtesten op cellen (*in vitro* bioassays), is gerelateerd aan een breed scala aan gemeten en gesimuleerde stofkenmerken en -eigenschappen die een rol spelen in de persistentie of mobiliteit van de stoffen in de waterketen. Het is belangrijk hierbij te vermelden dat de toepassing van data van testen op celniveau het proces van accumulatie, distributie, metabolisme en uitscheiding op organismeniveau niet meeneemt. De gemeten effecten representeren dus de toxiciteit op celniveau en moleculair niveau in een waterige matrix en niet de toxiciteit op organismeniveau (de mens). Het voorliggende onderzoek heeft voor 3360 stoffen gegevens verzameld van 534 toxiciteitstesten uit het ToxCast programma (U.S. EPA, 2015). Met behulp van twee statistische data-analysetechnieken is gekeken of er relaties konden worden gevonden tussen stofkenmerken en -eigenschappen die mobiliteit en persistentie beïnvloeden en de toxiciteit van de stof in de testen uitgedrukt in 'effectconcentraties', de concentratie waarbij een bepaald effect wordt gemeten. In het huidige onderzoek betreffen deze effectconcentraties AC_{50}s, wat staat voor de actieve concentraties waarop 50% van de biologische activiteit is waargenomen

Van de beschikbare stoffen en toxiciteitstesten zijn 111 tests voor 2845 stoffen geselecteerd voor verdere analyse. Daarbij zijn alleen testen met voldoende datapunten (geteste stoffen) en stoffen met betrouwbare gegevens geselecteerd. Bewust zijn zeer hydrofobe stoffen (met een octanol-water partitiecoëfficiënt van 100.000 of hoger) uit de data verwijderd, omdat het testen van dergelijke stoffen problematisch is, en gegevens van dergelijke studies minder betrouwbaar (S. E. Escher et al., 2022) omdat ze plakken aan de wand en alle organische fracties in het testsysteem en dus uit de waterige oplossing verdwijnen. De onderzochte toxiciteitstesten dekken een breed scala aan toxicologische mechanismen af, van specifieke receptor-gebaseerde hormonale effecten tot generieke DNA- en celschade.

Om relaties tussen stofeigenschappen en toxiciteit te analyseren zijn twee verschillende data-analysetechnieken gebruikt. (1) Met behulp van zogenoemde random forest analyse worden patronen in de data onderzocht zonder dat er lineaire verbanden hoeven te bestaan (Janitza & Hornung, 2018). Dit wordt gedaan door data via beslisbomen hiërarchisch te clusteren op basis van hun toxiciteit en te onderzoeken welke stofeigenschappen het meest bijdragen aan een goede voorspelling van de toxiciteit van de stoffen. (2) Uit alle beschikbare stofkenmerken of -eigenschappen (>1000) zijn dampdruk, molecuulgewicht, log octanol-water partitiecoëfficiënt, log organisch koolstof-water sorptiecoëfficiënt, kookpunt, en de berekende biodegradatiesnelheid als meest voorspellend gekenmerkt door dede random forest analyse. Deze stofeigenschappen zijn geselecteerd voor verdere analyse. Door middel van regressieanalyses zijn lineaire verbanden tussen stofeigenschappen en toxiciteit onderzocht, zoals is gedaan in Lambert et al. (2022), én zijn deze eigenschappen gerelateerd aan de mobiliteit en persistentie van de stoffen in het (water)milieu. Met beide methoden kunnen ook voorspellingen gedaan worden over de toxiciteit voor stoffen die niet in de dataset staan, mits er statistisch significante relaties zichtbaar zijn en de eigenschappen van de onbekende stof zich niet buiten het bereik van de stoffen in de dataset bevinden waar het model mee is ontwikkeld.

Zowel de random forest analyse als de lineaire regressie is in eerste instantie uitgevoerd voor de gehele dataset (111 toxiciteit testen, 2845 stoffen). Vervolgens is binnen de beschikbare data specifiek gekeken naar stoffen die effecten vertonen op twee toxicologische eindpunten van hormoonverstoring. Dit betreft de respons op de oestrogeen (vrouwelijk geslachtshormoon) receptor en de zogenoemde Pregnane X receptor die een belangrijke rol speelt in het

controleren van genexpressie voor het metabolisme. Daarnaast is specifiek gekeken naar twee subsets van stoffen, namelijk geneesmiddelen en een sub-selectie van PMT-stoffen (met octanol-water partitiecoëfficiënten onder de 1000). Geneesmiddelen zijn als groep geselecteerd omdat deze stoffen zijn ontwikkeld om een specifiek effect te hebben op biologische systemen met veel verschillende werkingsmechanismen en daarom mogelijk veel activiteit zou worden waargenomen in de toxciteitstesten. Met de subset van polaire stoffen is onderzocht of binnen het domein van meer polaire stoffen nog steeds een verband te ontwaren is tussen polariteit en effect.

1.4 Resultaten samengevat

De random forest en regressieanalyses lieten zien dat binnen de beschikbare data een verband zichtbaar was tussen toxiciteit en stofkenmerken als molecuulgewicht, dampdruk en kookpunt en eigenschappen als de wateroplosbaarheid, sorptiecoëfficiënten tussen water organisch materiaal in bodem (log K_{oc}) of partitiecoëfficiënten tussen water en octanol (log K_{ow}) en biodegradatie. Met name log K_{oc} en biodegradatie zijn sterk gerelateerd aan respectievelijk mobiliteit en persistentie. Stoffen worden gecategoriseerd als PM-stof wanneer de halfwaardetijd voor biologische afbreekbaarheid in zoet water (12 °C) meer dan 40 dagen bedraagt, en de log Koc van deze stof onder de 4 bedraagt. Deze parameters zijn derhalve verder onderzocht in de lineaire regressieanalyse. De nauwkeurigheid van het random forest model was echter relatief laag. Dit wordt veroorzaakt door ruis in de data en door de – soms relatief kleine steekproefgroottes in de modellen. Tevens kon geen statistisch onderscheid gemaakt worden tussen individuele toxiciteitstesten, daarom is de random forest analyse alleen uitgevoerd voor de dataset als geheel, met alle toxiciteitstesten en stoffen tezamen.

Stofeigenschappen en kenmerken in relatie tot persistentie en mobiliteit

De partitie tussen octanol en water en distributie tussen organisch materiaal en water zijn eigenschappen die sterk correleren met de mobiliteit van stoffen in watersystemen (Neumann & Schliebner, 2019; Reemtsma et al., 2016). De eigenschap biodegradeerbaarheid is een maat voor persistentie in systemen waar organismen zorgen voor de omzetting en/of mineralisatie van stoffen (Neumann & Schliebner, 2019). Het kookpunt en de dampdruk zijn stofkenmerken die een rol spelen in de verdeling tussen bodem of water en lucht en daarmee dus ook verband houden met de mobiliteit. Door de correlaties tussen toxiciteit en deze eigenschappen of kenmerken te bestuderen kan de vraag worden beantwoord of de karakteristieken die een stof tot een PMOC maken ze ook (intrinsiek) minder giftig maken.

De onderstaande heatmap (Tabel 1) vat de relaties en correlaties tussen stofeigenschappen en toxiciteit samen voor zowel de gehele set aan stoffen waarvoor relevante toxicologische informatie beschikbaar was, als voor specifieke subsets van stoffen. Relaties tussen de toxiciteit van de verbinding (uitgedrukt als AC50) en de fysisch-chemische eigenschappen, verkregen uit lineaire regressieanalyses. Significante correlaties worden weergegeven als blauwe (statistiek significante negatieve correlatie tussen de fysisch-chemische parameter en de AC_{50}) en groene (statistisch significante positieve correlatie tussen de fysisch-chemische parameter en de AC_{50}) cellen, terwijl grijze cellen de afwezigheid van significant lineaire verbanden aangeven. ER = oestrogeenreceptor; PXR = pregnane X-receptor; PMT = persistent, mobiel en toxisch. Hoewel gecorreleerde stofeigenschappen ($r > 0.97$) uit de dataset zijn verwijderd, zijn een eigenschappen toch aan elkaar gerelateerd. Zo is er bijvoorbeeld een omgekeerd evenredige relatie tussen dampdruk en kookpunt en zijn ook molecuulgewicht en octanol-water partitiecoëfficiënt. De reden om deze stofeigenschappen toch apart te beschouwen is de databeschikbaarheid voor deze parameters.

Tabel 1 Relaties tussen de toxiciteit van de verbinding (uitgedrukt als AC_{50}) en de fysisch-chemische eigenschappen, verkregen uit lineaire regressieanalyses. Significante correlaties worden weergegeven als blauwe (statistiek significante negatieve correlatie tussen de fysisch-chemische parameter en de AC_{50}) en groene (statistisch significante positieve correlatie tussen de fysisch-chemische parameter en de AC_{50}) cellen,

terwijl grijze cellen de afwezigheid van significant lineaire verbanden aangeven. ER = oestrogeenreceptor; PXR = pregnane X-receptor; PMT = persistent, mobiel en toxisch. De waarden in de tabel geven de p-waarde aan, wanneer de regressiecoëfficiënten worden vergeleken met 0.

Fysisch-chemische karakteristieken	**Alle stoffen**	**ER-actieve stoffen**	**PXR-actieve stoffen**	**PMT-stoffen**	**Genees-middelen**
log K_{oc}	P = 0.0001	P = 0.0001	-	P = 0.038	P = 0.017
log K_{ow}	P = 0.0002	P = 0.0348	P = 0.823	P = 0.021	P = 0.153
Biodegradatie halfwaardetijd	P = 0.0031	P = 0.0083	-	P = 0.362	P = 0.692
Dampdruk	P = 0.0462	P = 0.0772	P = 0.1033	P = 0.506	-
Kookpunt	P = 0.00048	P = 0.0004	-	P = 0.0096	P = 0.042
Molecuulgewicht	P = 0.871	P = 0.619	P = 0.0347	P = 0.0001	P = 0.748

De lineaire regressieanalyses van de gehele dataset lieten significant negatieve correlaties zien tussen toxiciteit en de log octanol-water partitiecoëfficiënt (log K_{ow}), de log organisch koolstof-water distributiecoëfficiënt (log K_{oc}), het kookpunt, het molecuulgewicht en de biodegradatiesnelheid (kolom 2 'Alle stoffen'). Dit wil zeggen dat een hogere waarde van een van deze parameters leidt tot een lagere effectconcentratie (AC_{50}), en dus een hogere (intrinsieke) toxiciteit. De dampdruk had een positieve correlatie met de effectconcentratie (AC_{50}), wat wil zeggen dat een hogere dampdruk leidt tot een hogere effectconcentratie en dus een lagere toxiciteit.

Bij het bestuderen van subsets van stoffen uit de dataset (hormoonverstorende stoffen, geneesmiddelen en polaire stoffen), werd bepaald of de in de gehele dataset waargenomen correlatie tussen toxiciteit en de geselecteerde stofeigenschappen standhield. In kolommen 3 'ER-actieve stoffen' tot 6 'Geneesmiddelen' is te zien dat de correlatie tussen log K_{oc} en toxiciteit in de gehele dataset gehandhaafd blijft voor de subsets van stoffen met specifieke effecten (ER-stoffen, PXR-stoffen), toepassingen (geneesmiddelen) én wanneer enkel de meer polaire persistente en toxische stoffen worden geselecteerd (PMT-stoffen). De relatie tussen log K_{oc} en toxiciteit blijft dus gehandhaafd, ongeacht de subset van stoffen. Voor indicator voor persistentie (biodegradatiehalfwaardetijd), aangezien enkel voor ER-actieve stoffen een significante correlatie waar te nemen is tussen halfwaardetijd en toxiciteit. Ook voor dampdruk, kookpunt, molecuulgewicht en de log K_{ow} bleek dit niet het geval.

Het handhaven van de correlatie voor log K_{oc}, wat een belangrijke parameter is die de mobiliteit van stoffen in water-sediment en water-bodemsystemen bepaalt, suggereert dat stoffen die mobieler zijn (met een lagere log K_{oc}) een hogere AC_{50} hebben, en dus intrinsiek minder toxisch zijn, ongeacht de selectie van stoffen. De waargenomen correlatie met biodegradatiesnelheid in de totale dataset verviel voor geneesmiddelen, PMT-stoffen en PXR-actieve stoffen. Dit doet vermoeden dat biodegradeerbaarheid geen of slechts een zwak direct verband houdt met (intrinsieke) toxiciteit. Mogelijk zijn de waargenomen significante correlaties voor ER-actieve stoffen en de gehele dataset het gevolg van een bekend artefact in de uitgevoerde toxiciteitstests, specifiek voor gemakkelijk afbreekbare stoffen. De AC_{50}-waarden in deze tests worden doorgaans bepaald op basis van nominale concentraties (concentraties zoals deze zijn toegevoegd aan het testsysteem) of gemeten initiële concentraties (gemeten concentraties van de stoffen bij aanvang van de experimenten). Voor stoffen die zeer makkelijk biologisch afbreken en dus een zeer korte halfwaardetijd hebben, is het aannemelijk dat de daadwerkelijke gemiddelde blootstellingsconcentratie aanmerkelijk lager is dan de veronderstelde initiële of gemeten startconcentratie. Daarmee wordt de toxiciteit van de snel biodegradeerbare stoffen dus onderschat, tenzij de afbraakproducten van deze stoffen ook toxisch zijn. Een mogelijke onderschatting van de toxiciteit van biologisch afbreekbare stoffen wordt ondersteund door de waarneming dat voor PMT-stoffen (die per definitie niet zeer snel biologisch afbreken) en geneesmiddelen (waarvoor biologische stabiliteit meestal een ontwerpcriterium is) geen significante correlatie wordt waargenomen. In de praktijk zullen snel biologisch afbreekbare stoffen echter waarschijnlijk ook afbreken of

omzetten tussen het laatst gemeten punt (afpomp) en het punt waarop inname in de mens plaatsvindt (de kraan). Aangezien drinkwater slechts sporadisch wordt bemonsterd, zal deze onderschatting van de toxiciteit in de praktijk hoogstwaarschijnlijk niet tot problemen leiden.

Zijn persistente mobiele stoffen minder giftig?

Op de vraag of PMOC (intrinsiek) minder giftig zijn kan dus het volgende antwoord worden gegeven. Kenmerken die de mobiliteit van de stoffen beïnvloeden zoals log K_{oc} correleren negatief met de AC_{50}, wat betekent dat deze stoffen gemiddeld genomen inderdaad minder giftig zijn. Terwijl voor stoffen die persistenter zijn, op basis van hun biodegradatiehalfwaardetijd, deze relatie niet duidelijk naar voren komt: enkel voor ER-actieve stoffen is immers een statistisch significante (positieve) correlatie aangetroffen tussen persistentie en toxiciteit. Hoewel de correlatie tussen log K_{oc} en toxiciteit significant is, kan maar een deel van de variatie binnen toxiciteitsgegevens door de log K_{oc} worden verklaard. Bovendien is het maar voor een beperkt aantal toxicologische mechanismen bestudeerd omdat de data niet toereikend was om dit voor andere mechanismen te bestuderen.

1.5 Van *in vitro* toxiciteit naar effecten op organismeniveau

In de bovenstaande beschrijvingen wordt gesproken over de intrinsieke toxiciteit van de stoffen. De term ‘intrinsiek is toegevoegd omdat de bepaling van de toxiciteit in de gebruikte testen gebaseerd is op de interactie van de stof met een specifieke receptor of celstructuur en een specifieke respons wordt gemeten. Deze specifieke respons is niet direct te vertalen naar het toxicologische effect op het niveau van het gehele organisme. Hoe dit zich vertaalt naar het toxicologische effect op het niveau van een heel organisme is complex en onderwerp van veel onderzoek (Zie bijvoorbeeld het BTO-rapport BTO.2018.030) (Bajard et al., 2023; Yoon et al., 2012). De vraag rijst of *in vitro* assays wel de juiste testen zijn om daadwerkelijke effecten op organismeniveau te bepalen. Wellicht dat deze assays te conservatief zijn of het effect juist onderschatten. Meer onderzoek is dan ook nodig om te bepalen of toxiciteitsgegevens op basis van *in vitro* assays wel geschikt zijn in de prioritering van (PMOC-)stoffen. Dit is ten eerste omdat de blootstelling in deze testen moet worden vertaald naar de blootstelling of dosis van het gehele organisme. Daarbij kan accumulatie, verdeling, metabolisme, en uitscheiding van de stof, de blootstelling op celniveau in het organisme sterk beïnvloeden (Breen et al., 2021). Ten tweede worden responsen op celniveau gemeten die verschillen van de uiteindelijk effecten op organismeniveau door diverse fysiologische reacties op de aanwezigheid van dergelijke stoffen. Ten derde wordt in de toxiciteitstesten enkel gekeken naar effect veroorzaakt door een individuele stof, terwijl organismen continu worden blootgesteld aan complexe mengsels van stoffen, waarin de toxicologische werking van stoffen kan worden beïnvloed door andere stoffen. Conclusie en relevantie voor de watersector

Wat betekenen deze bevindingen voor de drinkwatersector?

Uit de analyses blijkt dat de significante relatie tussen persistentie en toxiciteit voor de gehele dataset niet overeind blijft als stoffen met specifieke effecten (ER-stoffen), toepassingen (geneesmiddelen) of andere eigenschappen (PMT-stoffen) worden geselecteerd. Mogelijk is deze geobserveerde correlatie tussen persistentie en toxiciteit voor ER-stoffen een artefact, omdat zeer makkelijk biodegradeerbare stoffen moeilijk te toetsen zijn in een *in vitro* toxiciteitstest. Hoe dan ook is dit vanuit de praktijk minder relevant, want in de afvalwaterzuivering, het oppervlaktewater, het grondwater én de drinkwaterzuivering zijn talloze biodegradatieprocessen mogelijk die de uiteindelijke blootstelling van mens en milieu aan deze makkelijk afbreekbare stoffen beperken. Waarmee de vraag of ze al dan niet minder toxisch zijn dus ook minder urgent wordt. Makkelijk afbreekbare stoffen kunnen echter wel worden omgezet in persistente transformatieproducten, die wel relevant kunnen zijn. Want stoffen die zeer persistent zijn, zijn per definitie extra relevant, omdat hun persistentie maakt dat ze hoe-dan-ook lang in het (water)milieu aanwezig zullen blijven en daarin bij voortdurende emissies ook kunnen ophopen.

Uit de analyses blijkt wel dat er significante correlaties bestaan tussen de intrinsieke toxiciteit op celniveau en diverse kenmerken en eigenschappen van stoffen die hun mobiliteit in watersystemen, bodempassage, en waterbehandeling met sorbentia beïnvloeden. Dit kan wellicht verklaard worden doordat de in vitro assays vaak in water worden uitgevoerd (Fischer et al., 2017). Correlaties tussen intrinsieke toxiciteit en stofeigenschappen worden gezien voor de gehele dataset en voor subsets van stoffen (ER-actieve stoffen, PXR-actieve stoffen, geneesmiddelen en PMT-stoffen). Ze kunnen echter maar een beperkt deel van de variatie in de toxiciteit verklaren, wat betekent dat een polaire stof gemiddeld wel minder toxisch is maar dat niet elke polaire stof beperkt toxisch is. Daarmee zijn de gevonden correlaties dus ook niet direct toepasbaar in risicobeoordeling. De gevonden correlaties geven echter wel een indicatie dat persistentie evenredig schaalt met biologische activiteit en dat mobiliteit omgekeerd evenredig schaalt met biologische activiteit, wat van waarde kan zijn in een (eerste) prioritering van stoffen.
De hogere toxiciteit van bepaalde chemische stoffen in vergelijking met anderen houdt óók verband met bepaalde chemische structuren of kenmerken die leiden tot toxische activiteit via specifieke werkingsmechanismen, waarvan er in deze studie (maar) twee zijn behandeld. Dit veroorzaakt een grotere mate van ruis in de dataset, wat voor verstorende correlaties kan zorgen. Dit betekent dat nader onderzoek nodig is naar (verdere) classificatie van stoffen en de relatie tussen toxiciteit, mobiliteit en persistentie binnen klassen van stoffen met soortgelijke werkingsmechanismen of chemische structuren, zoals bijvoorbeeld bepaalde typen PFAS (Per- en polyfluoralkylstoffen). Hoewel classificatie van stoffen zorgt voor het verkleinen van datasets, kan – waar mogelijk - op deze wijze in meer detail worden bekeken hoe de variatie van mobiliteit van stoffen binnen een dergelijke klasse (een gedefinieerd chemisch domein) invloed heeft op de toxiciteit voor bepaalde eindpunten. Dit wordt in het vervolgonderzoek '*Deeper understanding of PMOC toxicity'* bestudeerd.

De significante inverse correlatie tussen mobiliteit en toxiciteit is enigszins geruststellend, maar uiteindelijk maakt zowel de toxiciteit als de blootstelling het risico (Kavlock et al., 2018). De M (en P) eigenschappen bevorderen de blootstelling via water doordat dergelijke stoffen simpelweg niet of nauwelijks worden afgebroken en vrijelijk met het water door het milieu en waterzuiveringsprocessen kunnen bewegen. Dit betekent dat het uiteindelijke risico, ondanks een eventuele beperkte intrinsieke toxiciteit, op termijn groter kan worden.. Hoewel PMOCs een relatief lagere kans tot toxische effecten lijken te hebben, is het om deze reden van belang om blootstelling aan deze stoffen regelmatig te evalueren. Het is momenteel, ook na dit onderzoek, nog onduidelijk of het risico van deze stoffen significant is.

Onlangs is een scoresysteem voor PMT-identificatie ontwikkeld waarin een combinatie van kwalitatieve en kwantitatieve informatie over fysisch-chemische eigenschappen en moleculaire structuren van chemische stoffen is verwerkt (Hartmann et al., 2022). Dit is een waardevolle stap naar de identificatie en classificatie van PMT-chemicaliën in databanken zoals REACH. Er moet echter nog een volledig kwantitatief model worden ontwikkeld, dat kan worden gebaseerd op de hier gepresenteerde resultaten. Het mechanistisch inzicht in de toxiciteit van PMOC's dat dergelijke modellen zullen opleveren, gecombineerd met gemeten toxiciteitsgegevens van high-throughput *in vitro* testen zoals in deze studie zijn gebruikt, maakt het mogelijk de kennisklоof inzake de toxiciteit van PMOC's te dichten. Bovendien zal dit ook het broodnodige inzicht verschaffen in additieve effecten van de mengsels van PMT's die alomtegenwoordig zijn in de huidige waterbronnen en -cycli (Neuwald et al. 2022). Samen kan dit bijdragen tot het toekomstige risicobeheer van PMT's dat nodig is om te voldoen aan de ambitie van de Europese Unie van een gifvrij milieu tegen het jaar 2050 (European Comission, 2020), ook geldend voor die chemische stoffen die het moeilijkst te verwijderen zijn uit waterige matrices.

2 Are PMOCs Less Toxic? Linking Physicochemical Compound Properties With Measured Toxicity

Renske P. J. Hoondert*[1], Milo L. de Baat*[1], Tessa E. Pronk[1], Milou M. L. Dingemans[1,2], Thomas L. ter Laak[1, 3]

* Both authors contributed equally to the manuscript.

[1] KWR Water Research Institute, P.O. Box 1072, 3430 BB, Nieuwegein, The Netherlands.

[2] Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands

[3] Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, Amsterdam 1098 XH, the Netherlands

Nb. This chapter is in preparation for publication

2.1 Abstract

Persistent and mobile organic chemicals (PMOCs) are gaining attention as a threat to the quality of water resources. Their high mobility (as a result of their polarity) and persistence lead to the presence and accumulation of PMOCs in surface water and drinking water sources. For many of these compounds, little is known about their toxicity due to the only recent attention for this aspect. As a result, health risk assessment of PMOCs is at an early stage. The present study is designed to gain insight into associations between toxicity of chemicals and their physicochemical properties that determine the persistence and mobility in the environment. The hypothesis that the mobility of PMOCs makes them less toxic than non-polar compounds was tested compiling a dataset matching physicochemical data for 3360 compounds with their measured effects and effect concentrations in 534 toxicity tests from the ToxCast program. Random forest analyses identified the physicochemical properties that relate most strongly to the induced effects and linear regression analyses quantified the strength and direction of these relationships. The analyses indeed showed that properties related to polarity, particularly K_{OW} and K_{OC}, are inversely related to effect concentrations, confirming that, in general, more polar compounds are less toxic. The division of the dataset into sub-datasets illustrated that these associations are differentially pronounced depending on the included subgroups of chemicals linked to modes of action, providing mechanistic insight into their toxicity. The associations presented here indicate that PMOCs interact less with tissues, cell membranes, and receptors than similar but more hydrophobic compounds, leading to lower intrinsic toxicity. More research is needed to allow more reliable predictions for specific substance groups. As the diversity and pervasiveness of PMOC in the water cycle lead to continuous exposure, further investigations into their potential threat to human health and water quality are warranted.

2.2 Introduction

In 2016, Reemtsma and colleagues put persistent and mobile organic chemicals (PMOCs) in the spotlight as a threat to the quality of water resources (Reemtsma et al., 2016). The high mobility (related to the polarity) of these chemicals allows them to pass through subsurface environments, evade sorption-based wastewater and in some cases even drinking water treatment processes. Their persistence means that (microbial) degradation processes have only a very limited effect on their concentrations throughout the environment. This leads to the likely accumulation of PMOCs in partially-closed water cycles (i.e. water cycles in which water extraction can occur downstream from the input of effluent and runoff) that are commonly used for drinking water production in urbanized areas (Neuwald et al., 2022; Schulze et al., 2019). This was corroborated in a study by Sjerps et al. (2016) which followed organic contaminant concentrations on their way from wastewater treatment plant (WWTP) effluent via surface- and groundwater to drinking water. While the total organic contaminant concentration decreased about 100-fold, the more polar contaminants (such as pharmaceuticals) remained in the water despite passage through the subsurface and multiple water treatment processes. Similarly, Gollong et al (2022) found a 270-fold reduction of non-polar features during wastewater treatment against only a 4-fold reduction of polar features, and a significantly higher occurrence of polar features in WWTP effluent and purified drinking water. This persistence of PMOCs in the water cycles used for drinking water production may very well lead to the frequent or even continuous (chronic) exposure of consumers (Schulze et al., 2018).

At the same time, Reemtsma et al. (2016) identified knowledge gaps related to the analysis, monitoring, water treatment, and regulation of PMOCs and concluded that these need to be closed to safeguard valuable water resources. Since then, important scientific progress has been made to narrow these knowledge gaps. For example, specialized analytical workflows have been developed that can expand the scope of regular chemical monitoring frameworks to include more hydrophilic neutral as well as permanently charged compounds (Montes et al., 2019; Schulze et al., 2019). Simultaneously, novel insights into the factors that govern compound removal allow the design of more effective water treatments for PMOCs (Albergamo et al., 2019; Zhou et al., 2021). Importantly, new policies to regulate and restrict persistent, mobile and toxic (PMT) and very persistent and very mobile (vPvM) substances, two subclasses of PMOCs, are currently well underway at a European level (Hale et al., 2022).

Nevertheless, formerly unidentified knowledge gaps regarding PMOCs have also become apparent. One of the widest knowledge gaps is the knowledge of PMT/vPvM and PMOC toxicity (Hale et al., 2022). Despite the fact that this vast group of compounds is widely dispersed in the environment, it only recently attracted wide attention in science and society, and little is known about the toxicities of these compounds. Additionally, there are analytical challenges to measure PMOCs. As a result, despite the growing body of environmental concentration data, risk assessment of PMOCs is still at an early stage. Particular attention has been given to the toxicity and mixture risk assessment of certain PMT compounds, in particular the large group of per- and polyfluoroalkyl substances (PFAS) (Bil et al., 2023). The research presented here, however, focuses on a fundamental broader and mechanistic understanding of the relationship between the physicochemical properties that define PMOCs and their human toxicity, which is yet to be established.

The correlation between the hydrophobicity of chemicals - commonly expressed as the log of the octanol/water partition coefficient (log K_{OW}) - and their toxic potency has been established for more than a century (Lambert et al., 2022). Generally, compounds with a higher log K_{OW}, and thus a higher hydrophobicity, have a higher toxic potency owing to their propensity to interact with and accumulate in cells and tissues (Verhaar et al., 1992). This was recently once again confirmed using a comprehensive database of 617 organic chemicals with curated and standardized acute toxicity data, showing that K_{OW} is inversely related to ecotoxicity (Lambert et al., 2022). Yet, the existence of PMT substances challenges the notion that hydrophilic chemicals are by definition of lower toxicological concern. This is partly attributable to the continuous presence of PMOCs in water cycles owing to their high mobility and persistence, leading to the chronic exposure of humans and the environment. Accordingly, Hale et al. (2020) postulated that PMTs pose an equivalent level of concern as persistent bioaccumulative and toxic compounds (PBTs). The European Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) program requires a PBT assessment for all registered substances. However, the equivalent does as of yet not exist for PMT/vPvM compounds, potentially causing chemical hazards to be overlooked during the registration process. Therefore, efforts are underway to make PMT and vPvM official hazard classifications under the Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures (CLP Regulation), making PMT assessment an integral part of a chemical safety assessment under REACH (European Commission, 2022b). To facilitate this, and given the knowledge gap on PMOC toxicity, an expansion of knowledge on compound toxicity to the polarity range that PMOCs fall into is warranted to allow a reliable assessment of chemical hazards.

To address this knowledge gap, the present study aimed to gain more insight into the toxicity of compounds in relation to the physicochemical properties that determine the persistence and mobility of chemicals in the environment. The high polarity (related to hydrophilicity) of PMOCs suggests that they interact less with tissues, cell membranes, and receptors than lipophilic compounds. Fewer and weaker interactions, a lower reactivity, implies fewer disruptions to cell integrity and functioning and thus lower toxicities. Therefore, it was hypothesized that the mobile (polar) properties of PMOCs make them inherently less toxic than non-polar compounds. To test this hypothesis, a dataset was compiled matching physicochemical data for compounds particularly relevant to the drinking water cycle with their measured effects in toxicity tests from the ToxCast program (Dix et al., 2007; U.S. EPA, 2015). Random forest analyses identified the physicochemical properties that relate most strongly to the induced effects and linear regression analyses quantified the strength and direction of these relationships.

The exploration of the use of *in vitro* assay outcomes as proxies for (human) toxicity (endpoints) fits the paradigm shift in toxicology from traditional time- and resource-intensive toxicological assessment methods to high throughput new approach methodologies (NAMs), while preventing ADME characteristics from obscuring associations between chemicals and in cell response (Krewski et al., 2010). It taps into a much larger wealth of available toxicological data than conventional methods would provide, allowing a more indiscriminate inclusion of many chemicals and endpoints for consideration to improve and support the health risk assessment of PMOCs.

2.3 Methods

2.3.1 Data acquisition and data clean-up

For the analyses presented here, a data set was compiled containing toxicological and physicochemical data for 3360 water relevant chemicals. A list of (drinking) water relevant chemicals provided by the Dutch Institute for Health and the Environment (RIVM), was collated with *in vitro* toxicological data from the ToxCast program (U.S. EPA, 2015) and physicochemical data from multiple sources as outlined below. Toxicological data included as response variables in the analyses described below consisted of continuous data, while data related to physicochemical descriptors consisted of both continuous data and binary data (0/1).

Toxicological data. The toxicological data used in the analyses presented here encompassed half-maximal activity concentrations (AC_{50}; in µM) for 534 unique toxicity test types retrieved from the US EPA ToxCast program (U.S. EPA, 2015). The data were log-transformed and normalized between 0 and 1 in order to compare magnitudes of effects between chemicals, toxicity tests, and test types.

Physicochemical data. Physicochemical data were obtained from the CompTox Dashboard of the US Environmental Protection Agency (Williams et al., 2017). These data were complemented with data from EPI Suite™, OPERA and PubChem, including both experimental data and parameters estimated based on e.g. the octanol-water partitioning coefficient, log K_{ow} (Mansouri et al., 2018; NCBI, 2022; US EPA, 2022).

The complete dataset was then cleaned up for analysis based on multiple criteria:

Chemicals (data rows) were removed for which data on their SMILES, functional groups, octanol-water partition coefficient (log K_{ow}), molecular weight and water solubility were not available.

Since low solubility and often coinciding high hydrophobicity of a compound frequently affects the actual exposure in a toxicity test (generally leading to underestimation of its effect) (Groothuis et al., 2015), poorly soluble compounds (i.e. with a solubility in mM below the corresponding AC_{50}) and/or a high log K_{ow} (> 5) were removed from the data set (Jonker & Van der Heijden, 2007).

Physicochemical descriptors (data columns) that did not show any variation in their data ($\sigma = 0$) across the list of chemicals were removed, as parameters lacking variability by definition have no added value in predicting the response variable (e.g. AC_{50}).

Strongly correlated physicochemical descriptors ($r > 0.97$) were removed from the dataset to avoid collinearity and prediction bias (Næs & Mevik, 2001). The values of pair-wise correlations were considered. If values of two physicochemical descriptors have a high correlation (higher than 0.97), the variable with the largest mean absolute correlation was removed. This was for instance the case for some descriptors estimated based on log K_{ow}, or for descriptors that yield the same value, regardless of data source (e.g. molecular weight and some topological parameters). The value of 0.97 was chosen based on a trade-off between keeping as many physicochemical descriptors as possible but removing those that are near-similar.

Finally, all *in vitro* toxicity tests for which less than 30 data entries (tested chemicals) were available were disregarded, to ensure an unbiased modelling practice. This cut-off of 30 data entries was based on the Central Limit Theorem, where assumptions about the population distribution are meaningless if the sample size exceeds 30, since the sampling distribution approximates the standard normal distribution (Kwak & Kim, 2017).

These criteria resulted in a list of 111 toxicity tests that were analyzed separately, covering 2845 unique chemicals.

2.3.2 Random forest and linear regression analyses

Random forest is a supervised learning algorithm using an ensemble of decision trees, capable of performing both regression and classification tasks. The algorithm continually randomly selects a subset of physicochemical descriptors and subdivides the data based on these descriptors until a full tree is developed and analyzed for predictive power using the physicochemical descriptors. The algorithm arrives at the best explanatory properties by always prioritizing the decision trees with the properties over all micropollutants that perform best to explain toxicity. The randomization process reduces bias and decreases variance between and within trees. Random forest is, aside from its ability to build accurate classifiers, a good and non-parametric method for feature selection. To get more insight into PMOC toxicity and the underlying toxicological mechanisms, random forest analyses were performed based on a large set of physicochemical descriptors taking toxicological endpoints as a response variable for each toxicity test individually.

A fixed number of 5000 decision trees was used in the random forest analyses and the top 10 descriptors explaining the most variance in AC_{50} were reported. Random forest analysis does not require the response variable and/or the predictors to be normally distributed. However, linear regression, applied to quantify the magnitude and direction of the observed effect, assumes the data to be normally distributed. Therefore, physicochemical descriptors used as parameters in the random forest analyses were tested for normality and were transformed to obtain normal distribution. Parameters that showed to be non-normally distributed were not removed, since binary variables (0 = FALSE, 1 = TRUE), such as the presence or absence of certain structural alerts are by definition not normally distributed. A disadvantage of any machine learning model, including random forest, is that it is so complicated that it can only be applied as a computer model. It is therefore not intuitively easy to interpret. Additionally, the random forest output does not include quantitative regression coefficients and therefore does not provide insight into the magnitude or direction of the observed effect. Therefore, additional single linear regression analyses were performed to provide insight into the magnitude and direction of the relationship.

Multiple linear regression analyses were performed with the parameters that showed to be the strongest explanatory variables for the toxicity test response (AC_{50}) in the random forest analyses. Additionally, single linear regression analyses were performed for parameters that showed to be the strongest explanatory variables in the random forest analysis. This resulted in a regression coefficient (yielding information on the direction and magnitude of the effect) and a p-value (yielding information on the statistical significance of the reported effect). P-values below 0.05 were considered significant.

2.3.3 Model validation

The goodness-of-fit for the single linear regression models was determined based on the adjusted R^2 for each multiple regression model. The adjusted R^2 accounts for the number of explanatory variables included in the model. The results of the random forest analyses were validated by computing the root-mean-squared error (RMSE) for the training set and out-of-bag cases in each model, to summarize both random error and systematic bias. The RMSE for the out-of-bag cases (OOB) was calculated by randomly separating data in a training (2/3 of all data) set and validation or out-of-bag (OOB) set (1/3 of all data), according to Brieuc et al. (2018). Unexplained variance in toxicity by the random forest model would be attributable to either true random behavior of the (toxicity test) response or lack of fit of the model (when including solely physicochemical descriptors as explanatory variables). The RMSE was computed by geometrically averaging the squared differences between log-transformed observed or data in the training set (O) and predicted or data in the test (OOB) set (P) values for each of *n* data points (i) (Equation 1).

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (O_i - P_i)^2} \quad [1]$$

Additionally, the variance explained by the random forest model was reported.

2.3.4 Dividing by substance groups

Compounds with different specific toxic modes of action (MoA) are expected to elicit different responses in certain types of toxicity tests, i.e. endpoints. Since the total dataset is comprised of a large variety of toxicity tests and chemicals yielding different kinds of MoAs, the data were divided into four specific datasets (including two groups based on MoA); ER-active compounds, PXR-active compounds, PMT compounds and pharmaceuticals. These groups were chosen based on data availability and were analyzed separately using random forest analyses and linear regression analyses. The MoAs of compounds in two of these groups of compounds were hormone receptor agonism, represented by active binding to the estrogen receptor (ER-active) and induction of xenobiotic metabolism, represented as active binding to the pregnane X receptor (PXR-active). Lists with chemicals that are pharmaceuticals, PMTs and ER-active compounds were extracted from the US EPA CompTox Chemicals Dashboard (**https://comptox.epa.gov/dashboard/chemical-lists/**), matched with the present dataset, and combined with the physicochemical data, resulting in subsets for pharmaceuticals (n = 207, out of 953 compounds on the initial list with water relevant compounds), PMTs (n = 492, out of 1138 compounds on the initial list) and ER-active compounds (n = 322, out of 470 compounds on the initial list). The list of PXR-active chemicals (n = 97, out of 108 compounds on the initial list) was compiled by selecting chemicals in the present dataset with primary agonistic activity in PXR-related endpoints (See the Supplementary Information for the complete lists of compounds that have been included or excluded in the random forest analyses per substance group).

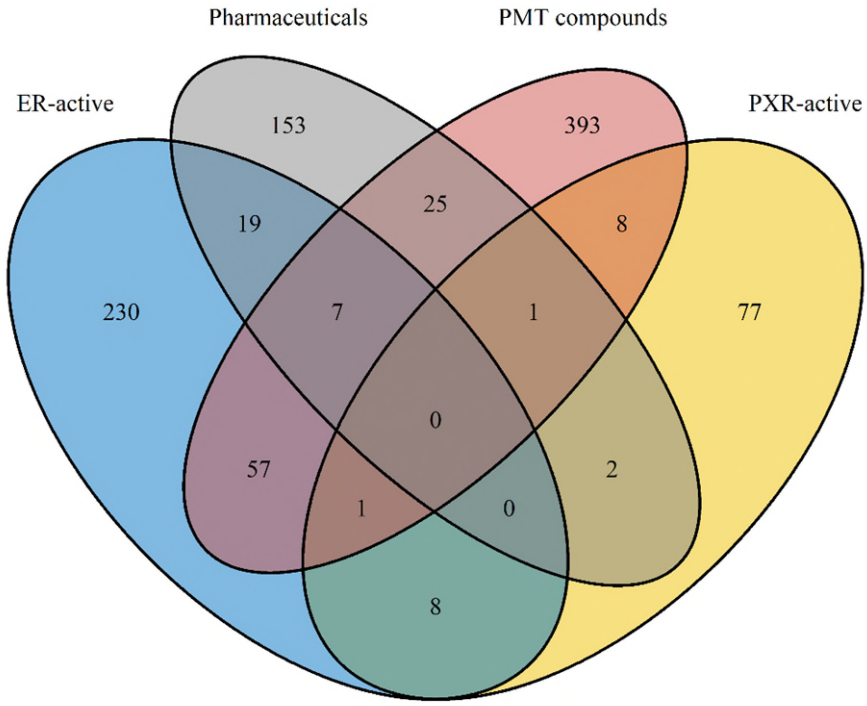


Figure 1: Venn diagram depicting overlap between compounds in the four subsets.

The four subsets of chemicals (pharmaceuticals and PMT compounds, and PXR-active and ER-active compounds) that were included in the random forest analyses and regression analyses were checked for overlap (Figure 1). Although the majorities of the subsets consist of unique CAS-numbers, there is overlap between pharmaceuticals and ER-active compounds (n = 19), PMT compounds and ER-active compounds (n = 57), and PMT-compounds and pharmaceuticals (n = 25).

2.4 Results & Discussion

2.4.1 Random forest analyses; linking chemical characteristics to toxicity

Out of the 534 toxicity tests included in the initial dataset, for a total of 111, the most important predictors (physicochemical and structural characteristics of chemicals) of toxicity were obtained for each toxicity test endpoint individually by using a random forest approach. The 10 most important predictors of toxicity (i.e. AC_{50}) for each assay were collected. The most frequently reported descriptors associated with responses over all assays were related to log K_{oc} (either experimental, based on K_{ow}, or predicted using the Molecular Connectivity Index method (MCI) (U.S. EPA, 2012)), log K_{ow}, vapor pressure, biodegradation, and boiling point (Figure 2). This implies that compound toxicity is mostly associated with volatility (e.g. vapor pressure and boiling point), persistence (biodegradation), chemical partitioning and distribution (related to log K_{ow} and log K_{oc}) and mobility (related to log K_{oc}). Additionally, structural complexity of the compound (both the elements contained and the displayed structural features including symmetry) is shown as a strong predictor of toxicity (PMComplexity), with 44 random forest analyses showing this topological

predictor in its top 10 strongest predictors. This parameter incorporates the number of elements as well as structural features of the chemical, such as symmetry.

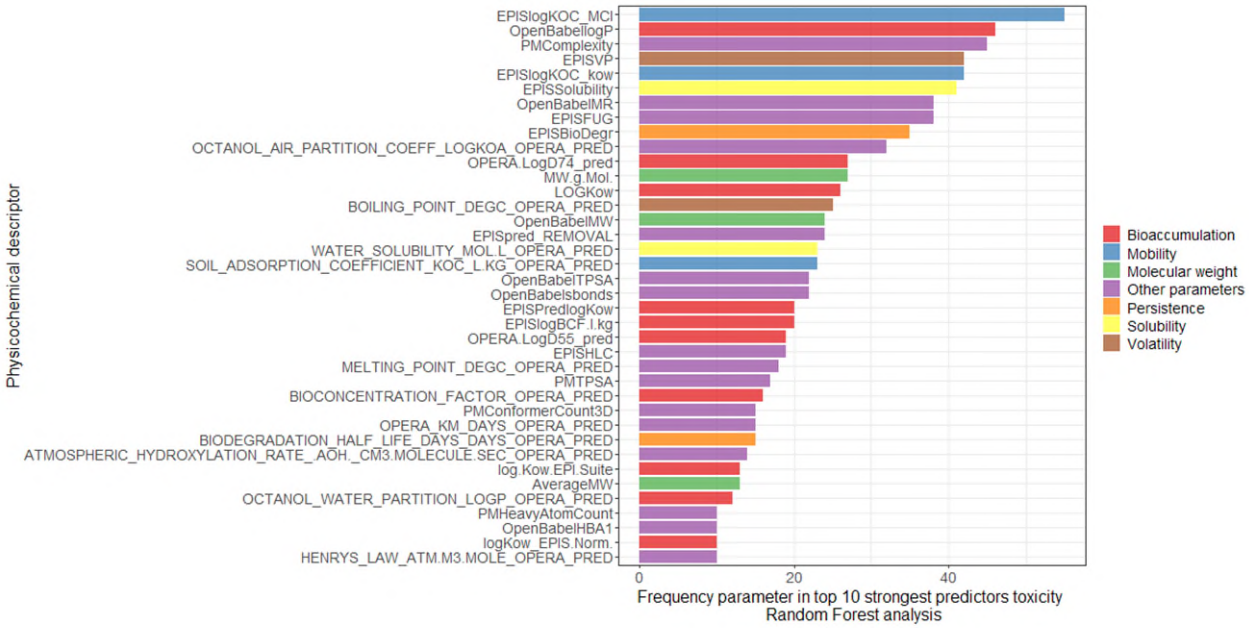


Figure 2: Frequency of physicochemical descriptors appearing in the top 10 descriptors with the highest predictive power after performing random forest analyses for 111 toxicity tests, excluding physicochemical descriptors with a frequency lower than 10.

Various sets of physicochemical descriptors were identified as the strongest predictors of toxicity for the different endpoints. The highest average variance explained by the random forest analyses was just over 50% for toxicity tests falling within the “miscellaneous” category (52.9% ± 21.6) followed by toxicity tests falling within the “developmental toxicity” category (50.73% ± 27.54) (Table 1). For all endpoints, the mean RMSEs for Out-of-bag samples were higher than RMSEs for the training set, which means that the accuracy of the random forest model was relatively low. This means that, in addition to a low explained variance, the predictive capacity of the random forest model was low, which is likely due to the limited number of chemicals that are available at this moment within a highly variable group of chemicals for each random forest analysis and for each individual toxicity test. A complete overview of variances explained, and RMSEs for all individual toxicity tests can be found in the Supplementary Information (Figures S3 and S4), which give insight into the predictive power of the random forest analyses and linear regression analyses, respectively.

Table 2: Number of toxicity tests, variance explained by random forest analyses (± standard deviation), adjusted R^2 when taking the six most commonly reported physicochemical descriptors, root-mean-square-error (RMSE) (± standard deviation), RMSE for out-of-bag cases (OOB) (± standard deviation) and most frequently reported strongest predictor in the random forest analysis for the seven most commonly analyzed endpoints in ToxCast. OPERA = OPERA model, EPI = Epi Suite ™ model, S.D. = standard deviation, HLC = half-life coefficient, MCI = Molecular connectivity indices, AR = androgen receptor, ER = estrogen receptor.

Endpoint	*Number of toxicity tests*	*Mean* (± S.D.) number of chemicals	*Variance explained by random forest* (± S.D.)	R^2 (± S.D.)	*RMSE* (± S.D.)	*RMSE-OOB* (± S.D.)	*Most frequent strong predictors of in vitro response*
Cell viability	26	653 ± 507	24.4 ± 20	0.08 ± 0.1	4.3 ± 0.4	15.4 ± 5.4	Boiling point (OPERA), Biodegradation rate (EPI)
Development	10	148 ± 92	50.7 ± 27.5	0.18 ± 0.17	4.4 ± 0.6	14.0 ± 5.9	BCF (EPI), Log Koc (based on Kow) (EPI)
DNA damage	12	209 ± 152	38.1 ± 17.4	0.13 ± 0.09	4.7 ± 0.6	13.6 ± 6.4	Vapor pressure (EPI), HLC (EPI)
Endocrine (AR)	8	19 ± 26	42.5 ± 15.5	0.17 ± 0.11	4.1 ± 0.3	12.8 ± 2.7	Complexity (PubChem), Atmospheric hydroxylation rate (OPERA)
Endocrine (ER)	10	356 ± 262	43.4 ± 17.4	0.15 ± 0.15	4.6 ± 0.6	14.1 ± 4.9	Fugacity (EPI), Biodegradation rate (half-life) (CompTox)
Xenobiotic metabolism	9	153 ± 57	35.3 ± 18.8	0.13 ± 0.14	3.3 ± 0.6	10.3. ± 3.1	Density (CompTox), Log Koc (based on MCI) (EPI)
Miscellaneous	7	230 ± 32	52.9 ± 21.6	0.15 ± 0.2	4.0 ± 1.2	12.3 ± 2.6	Biodegradation rate (EPI), HLC (EPI)

2.4.2 Linear regression analysis; searching for linear relations between descriptors and response

Based on the total dataset, R^2s, illustrating the predictive power of physicochemical descriptors in the linear or multiple regression analyses with *in vitro* toxicity as a response variable, were found low compared to the variance explained by the random forest analyses for the same toxicity tests. This indicates that there probably is no linear relationship between physicochemical descriptors and the response variable as the modelling exercise was based on single linear responses only, disregarding any interaction between parameters or non-linear responses. The linear regression coefficients for the 111 analyzed *in vitro* assays, taking normalized AC_{50} values as response variables, were negative for boiling point (C°), log K_{oc}, log K_{ow}, biodegradation rate (half-life in days), and molecular weight (g/mol), with median regression coefficients of -0.074, -3.83, -0.29, -0.07, and -0.03 respectively, based on the absolute values

(Figure 2, upper graph). The median regression coefficients for log K_{ow}, log K_{oc}, boiling point and biodegradation rate all significantly differed from zero (p-value < 0.05; One sample t-test). Only the median value of regression coefficients for vapor pressure was positive (0.04), however median values for both vapor pressure (mmHg) and molecular weight did not differ significantly from zero (p > 0.05; One sample t-test). Hence, in general, the majority of the investigated physicochemical properties were inversely related to toxicity (expressed as AC_{50}, with a higher AC_{50} indicating a lower toxicity and a lower K_{oc}/K_{ow} indicating a higher mobility), albeit to varying degrees.

When standardizing the physicochemical descriptors by subtracting the values by the mean value per toxicity test and dividing the result by the standard deviation, rescaling the data to have a mean of zero and a standard deviation of one, we see a similar pattern, albeit more spread out (Figure S1). The influence of physicochemical descriptors that normally cover a large range of values (such as molecular weight, boiling point and biodegradation rate) become more apparent, as the values on the y-axis now inform us about the change in response (toxicity) when increasing the physicochemical descriptor by one standard deviation (i.e. a relative increase in descriptor value, rather than an absolute increase in descriptor value).

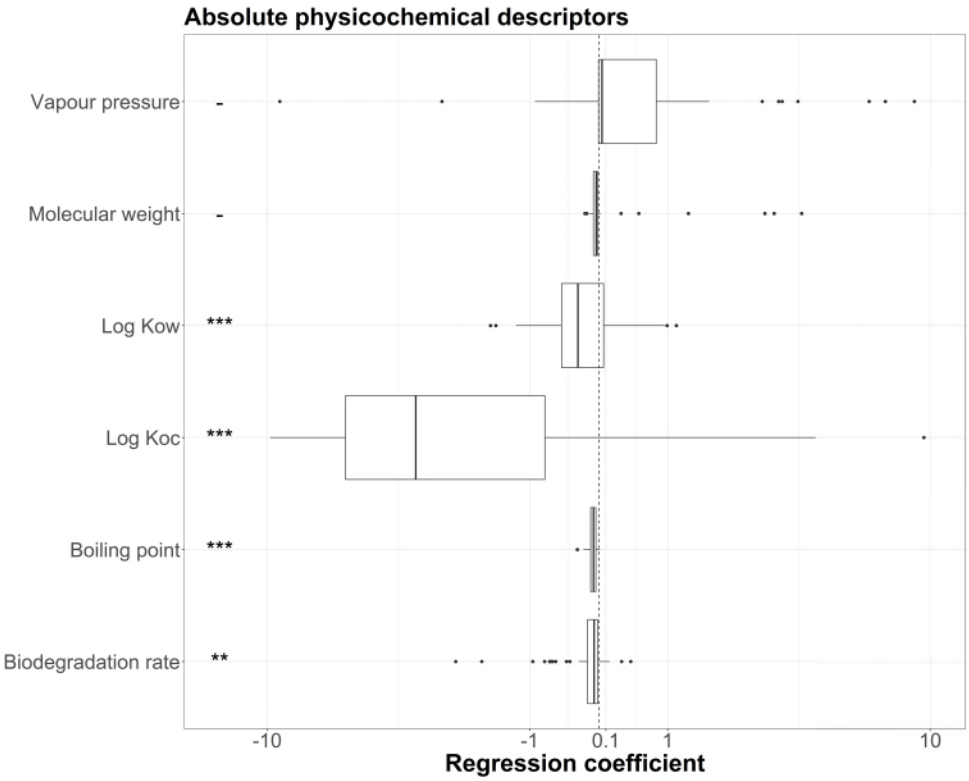


*Figure 3: Boxplots depicting the distribution of regression coefficients (minimum value, 25th percentile, median value, 75th percentile and maximum value) for vapor pressure (in mmHg), molecular weight (in g/mol), log octanol-water partition coefficient (log K_{ow}), log sorption coefficient to organic carbon (log K_{oc}), boiling point (in °C), and biodegradation rate (half-life in days) resulting from the random forest analysis for 111 bioassays. ***; $\mu \neq 0$, $p < 0.001$, **; $\mu \neq 0$, $p < 0.01$, *; $\mu \neq 0$, $p < 0.05$, -; $\mu = 0$, $p > 0.05$. Parameters with significantly similar distributions of regression coefficients were assigned a similar letter.*

These results indicate that, in general, AC_{50} activity concentrations decrease with a higher boiling point, biodegradation half-life, log octanol-water partition coefficient (log K_{ow}) and log sorption coefficient to organic carbon (log K_{oc}). This implies that compounds with a higher log K_{ow}, log K_{oc}, half-life or boiling point are generally more likely to evoke a response in a toxicity test and can thus be considered more toxic.

Mobility

Higher log K_{ow} values are associated with elevated acute toxicity due to their relevance to bioconcentration in membranes (Droge, 2018; B. Escher & Schwarzenbach, 1995). As a result, K_{OW} is a commonly applied descriptor in many aquatic toxicity models and has been reported as an important predictor of toxicity since the 1980s (Hutchinson et al., 1980; Lambert et al., 2022). Hence, the present findings, based on *in vitro* data, are in accordance with this long-established relationship between toxic potency and hydrophobicity expressed through the log K_{ow}. However, based on the present dataset, log K_{oc} appears to be the strongest descriptor of toxicity in many of the assessed toxicity tests (Figure 1). Interactions with organic carbon extend beyond hydrophobic interactions as – for example - electrostatic interactions and complexation of oppositely charged sorbents and sorbates lead to high sorption coefficients that is not reflected in octanol water partitioning (Hammer, 2019). This may suggest that those interactions with organic materials, might also correlate with interactions with biological structures and receptors leading to toxic effects. Such effects can be observed in *in vitro* assays as these generally do not include a substantial carbon-rich sink for chemicals, besides the cell model. How *in vitro* effect concentrations translate to real-world situations where sorption processes, among others, play an important role in the bioavailability, and thus exposure of organisms in the environment or target tissues within organisms, is an ongoing field of research (S. E. Escher et al., 2022; Hermens et al., 2007). Interestingly, K_{OC} is as of yet mainly considered for its role in mobility, while the present results reveal that chemical properties determining K_{OC} may also play an important role in the intrinsic toxicity of chemicals measured *in vitro*.

Persistence

Compounds are considered persistent based on their (bio)degradation rates, in which compounds with higher half-lives ($t_{1/2}$) are labelled as more persistent. The persistence of chemicals can lead to longer and/or more frequent exposure. Inversely, compounds that are readily (bio)degraded either within the environment or inside the organism have only a short time to exert their toxicity on organisms, tissues or cells. Persistency is therefore considered an important parameter in chemical hazard and risk assessment (Arp & Hale, 2022). The present results confirm its minor relevance in toxicity testing. This can be explained by the exposure within these *in vitro* assays. The timeframe of the majority of toxicity tests used in ToxCast falls in the range from less than an hour to 24 hours, with several exceptions that extend up to 6 days. These toxicity tests are biologically active systems, and often make use of nominal concentrations or at best measured concentrations at the start of the experiment. Biodegradation potential is probably indicative of how easily compounds are metabolized and inactivated by cells during *in vitro* assays. This leads to lower integrated exposure concentrations than initially added, thereby explaining the generally lower presumed toxicity of readily biodegradable chemicals in the present dataset (Jahnke et al., 2016).

Volatility

The boiling point of chemicals has not been described as a predictor of toxicity in the recent scientific literature. Its presently observed influence seems minor, but significant nonetheless. The explanation for this observation may be similar to that of the biodegradation potential, as very volatile chemicals can escape *in vitro* test systems. These test systems generally have small volumes and large surface areas exposed to air at room temperature or higher (>30°C for mammalian cell lines). Therefore, the most volatile chemicals might partially escape from the test matrix, leading to lower than nominal integrated exposure concentrations, explaining the generally lower presumed toxicity of the most volatile chemicals in the present dataset (Bakand & Hayes, 2010; Becker & Crass, 1982; Proença et al., 2021). The influence of volatility (described by boiling point) on toxicity test response in this case was more significant than the influence of chemical persistence. This may be due to non-normality of the biodegradation half lives used as input for persistence in the regression analyses.

2.4.3 Case studies of chemical groups and modes of action

Random forest models are distribution-free hierarchy-based model based on decision trees. Hence they do not rely on linear relationships between the predictor variables (physicochemical and structural descriptors of compounds) and the response variable (AC_{50}). The high RMSEs for out-of-bag cases compared to RMSEs for the full training dataset (Table 1) indicate that the random forest models constructed in the analyses are overfitted, and are probably not suitable for predicting toxicity of chemicals outside the training set. In many cases, the model performs worse than a random prediction in the absence of any associations between predictors and the response variable, which might be due to random noise (Janitza & Hornung, 2018). To counter this random noise, and thus decrease the differences between RMSEs for the training and test set, the sample size of the data acting as input in the random forest model needs to be sufficiently large (Mitchell, 2011). However, a statistically significant trend between the sample size of bioassays and the corresponding RMSEs ($p > 0.05$, $R^2 < 0.05$) was not observed in the present study. This implies that a larger sample size would not necessarily result in lower RMSEs, implying a better fit of the model. Using a more homogenic dataset of endpoints of a specific toxicological mode of action and/or a more homogenic set of chemicals may, however, reduce random noise and increase significance of the random forest analysis and improve linear correlations, improve predictions outside the training set and support mechanistic understanding.

The (cor)relations for subsets of data sorted according to mode of action (ER and PXR active compounds), PMT compounds and type of application (pharmaceuticals) are discussed in detail below, and are summarized in Table 2. These four subsets were analyzed separately in random forest analyses resulting in the explained variances for the available toxicity endpoints for each subset (Figure 4). Linear regressions allowed the quantification of the relationships between the chemical properties of the subsets with their *in vitro* toxicities (AC_{50}; Figures S5-8). The findings for each of the chemical subsets are discussed below and shown in Figure 5. Additionally, physicochemical descriptors used as predictor variables were standardized by subtracting the values by the mean value per toxicity test and dividing the result by the standard deviation, rescaling the data to have a mean of zero and a standard deviation of one. These results are shown in Figure S2 in the Supporting information.

Table 3: Relationships between compound toxicity (expressed as AC50) and physicochemical properties, obtained from linear regression analyses. Significant correlations are represented as blue (negative) and red (positive) cells, while grey cells indicate the absence of linear relationships. ER = estrogen receptor; PXR = pregnane X receptor; PMT = persistent, mobile and toxic.

Physicochemical properties	All compounds	ER-active	PXR-active	PMT	Pharmaceuticals
log K_{oc}					
log K_{ow}					
Biodegradation half-life					
Vapor pressure					
Boiling point					
Molecular weight					

Summarizing the results of the linear regression analyses, log K_{OC}, log K_{OW}, biodegradation rate, boiling point, and molecular weight all showed significant negative correlations with toxicity (expressed as AC50) for the entire dataset (Table 2). This means that a higher value of any of these descriptors correlates with a lower AC_{50}, and thus a higher (intrinsic) toxicity of the compounds. Only vapor pressure was significantly positively related to AC_{50}, implying that a higher vapor pressure relates to lower toxicity. Interestingly, when selecting various subclasses of compounds, a mixed picture of correlations between physicochemical properties and toxicity arises. The positive correlation between vapor pressure and toxicity is lost for all substance classes, while the correlation with boiling point is lost for PXR-active compounds and with molecular weight for ER-active compounds. Conversely, the positive correlation for log K_{OW} and biodegradation half-life are only maintained for ER-active compounds. Hence, the only descriptor that shows a stable significant correlation with toxicity, even when dividing into specific compound classes, is log K_{OC}. Log K_{OC} is closely related to the mobility of chemicals in soil and sediment-water systems.

ER-active compounds

When taking into account all endpoints available for the ER-active compounds, extracted from the US EPA CompTox Chemicals Dashboard, including all physicochemical descriptors in the random forest analysis resulted in an explained variance of 33.2%, based on 3541 individual data points for 322 chemicals. The ER-active compounds showed a high explained variance (>50%) for most of the available endpoints (Figure 4). However, in the present study we did account for the fact that endpoints may be mechanistically or biologically related. Additional research could focus on these correlations between endpoints according to adverse outcome pathways. The only endpoints for which a low explained variance was reported, based on the included physicochemical descriptors, were related to oxidative stress, metabolic activity and DNA damage. The strongest three predictors of toxicity identified by the random forest analysis were related to molecular weight, mobility (log K_{oc}/ Log K_{ow}) and vapor pressure. Mobility (log K_{oc}) showed to be strongly negatively associated with *in vitro* toxicity ($p < 0.001$), while no such association was observed between molecular weight and vapor pressure (Figure 5 and Figure S5). This indicates that chemical characteristics that affect binding to organic carbon may also affect ER-receptor binding, although this relationship is not thoroughly described in literature. The absence of the correlations with molecular weight might be associated to the size and dimensions of the molecules, as it is known that there is an optimal molecular size and shape for ER-receptor binding. Lower ER-activity is expected for either too small or too bulky chemicals (Hong et al., 2002). Furthermore a weak, albeit significant ($p < 0.05$), negative relation between toxicity (AC_{50}) and persistence (biodegradation) and boiling point could be observed for this specific subset of compounds. These observations are considered to be not plausible and can be related to experimental artifacts associated with the determination and control of exposure concentrations in applied tests systems (see also section *Are PMOCs less toxic?*).

PXR-active compounds

When taking into account all endpoints available for the PXR-active compounds, the random forest analysis resulted in an explained variance of 52.1%, based on 772 data points, covering 97 substances. These data were compiled by selecting chemicals in the present dataset with primary agonistic activity in PXR-related endpoints, including all physical-chemical descriptors. The PXR-active compounds showed a high explained variance (>50%) for most of the available endpoints (Figure 4). The only endpoints for which a low variance explained (<50%) was reported, based on the included physicochemical descriptors were related to DNA damage and (embryonic) development. The three strongest predictors identified by the random forest analysis were related to hydrophobicity (log P and log K_{ow}) and vapor pressure. Although only 9 PXR-active compounds overlap with ER-active compounds (Figure 1), associations between physicochemical descriptors and toxicity test response were very similar (Figure S5 and S6). However, none of the correlations between toxicity and indicators for persistence and mobility were significant (Figure 5 and Figure S6). This observation remains to be explained but might be affected by the relatively small dataset and/or to the promiscuous nature of the PXR. Compared to other nuclear receptors, the hydrophobic ligand-binding domain of the PXR is large and flexible, and contains polar residues, allowing it to bind a wide variety of ligands spanning a range of polarities (Watkins et al., 2001; Wu et al., 2013). This is a plausible explanation for the lack of a relationship between the polar properties of chemicals and the activation of the PXR.

PMT compounds

The class of PMT compounds (*See Dividing by substance groups*) showed a low explained variance (<50%) for most of the available endpoints (Figure 4). The only endpoints for which a high variance explained (> 50%) was reported, based on the included physicochemical descriptors were related to metabolic activity and specific receptors (aryl-hydrocarbon-, progesterone-, and androgen receptors). Nevertheless, taking into account all endpoints, including all physicochemical descriptors in the random forest analysis resulted in an explained variance of 62.7%, based on 4336 data points, covering 492 compounds. The three strongest predictors identified by the random forest analysis were related to the number of hydrogen bond acceptors 1 (HBA1), the number of bonds, and log K_{oc}, indicating that even within the limited K_{ow} range of this group of chemicals the relation of mobility to toxicity is apparent. A weak inverse relation was observed between *in vitro* AC_{50} and molecular weight, implying that larger molecules are more toxic. Additionally, a significant inverse relationship ($p < 0.05$) between AC_{50} and boiling point ($p < 0.01$) was observed for this large group of chemicals (Figure 5 and Figure S7), indicating that compounds with a higher boiling point are likely to be more toxic. This is probably partly due to the fact that these compounds are less likely to evaporate from the toxicity test well and more likely to stay in the medium, increasing exposure to the compound. As also observed for ER-active compounds, mobility (log K_{oc}) was significantly negatively correlated with AC_{50} values ($p < 0.05$) but no correlation ($p > 0.05$) was observed between AC_{50} values and persistence (biodegradation). This illustrates that despite narrowing the mobility window of the substances by selecting more hydrophilic chemicals, the relation with log K_{OC}, observed for the full dataset is maintained, while with losing the most easily degradable substances the correlation of this variable with toxicity is lost.

Pharmaceuticals

When taking into account all endpoints available for the pharmaceuticals, including all physicochemical descriptors in the random forest analysis resulted in an explained variance of 56.5%, based on 2487 data points, covering 207 compounds. The pharmaceuticals also showed a high explained variance (>50%) for most of the available endpoints (Figure 4) except metabolic activity and immune response (<50%). The three strongest predictors identified by the random forest analysis were related to persistence (biodegradation), topological polar surface area (TPSA) and hydrophobicity (log K_{ow}). While log K_{oc} showed to be significantly negatively correlated with *in vitro* toxicity ($p < 0.05$), no such correlation was observed for persistence (biodegradation) (Figure 5 and Figure S8). A weak, albeit significant ($p < 0.05$), relation between toxicity and molecular weight and boiling point could be observed for the pharmaceuticals. The relations of pharmaceuticals and PMT substances are somewhat similar. This makes sense since the intended properties of pharmaceuticals partially overlap with PMT properties. Pharmaceuticals need to be somewhat persistent as ready biodegradation compromises their ability to maintain the often required constant

concentrations in users. Moreover, pharmaceuticals are commonly designed to be bioactive, which at higher doses may be exhibited as toxicity.

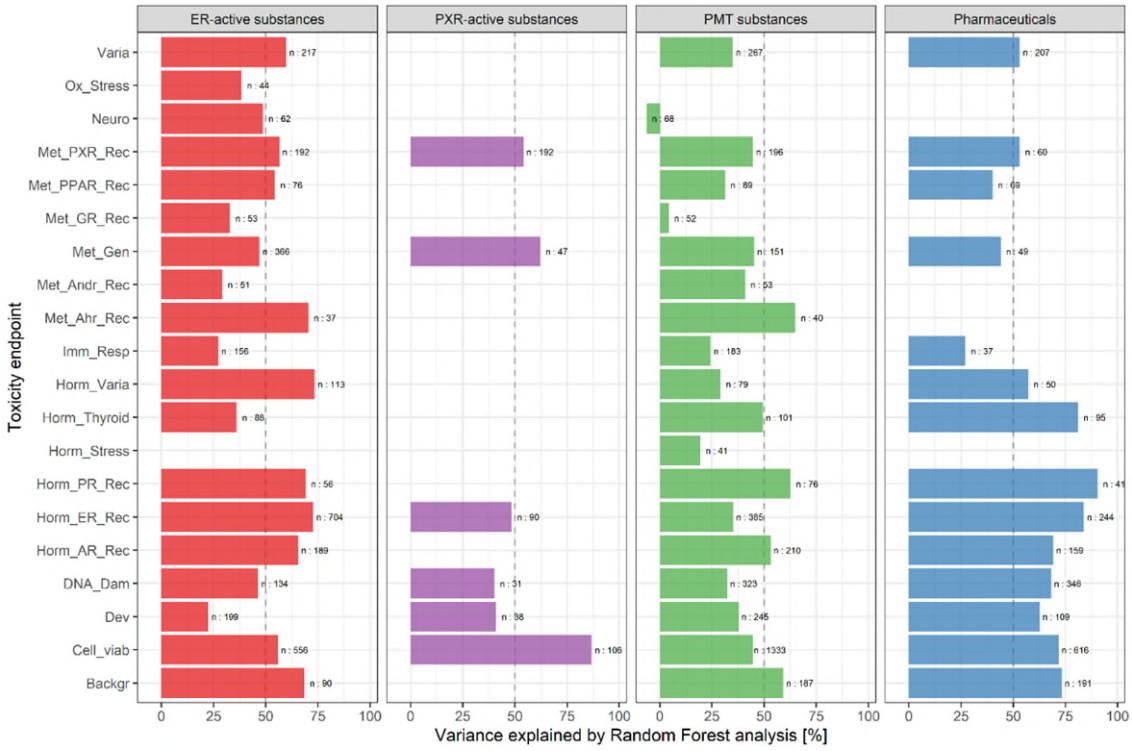


Figure 4: Variance explained by random forest analyses (in %) for each of four subsets of chemicals (ER-active compounds, Pharmaceuticals, PMT substances and PXR-active substances) for each group of toxicity endpoints from ToxCast.

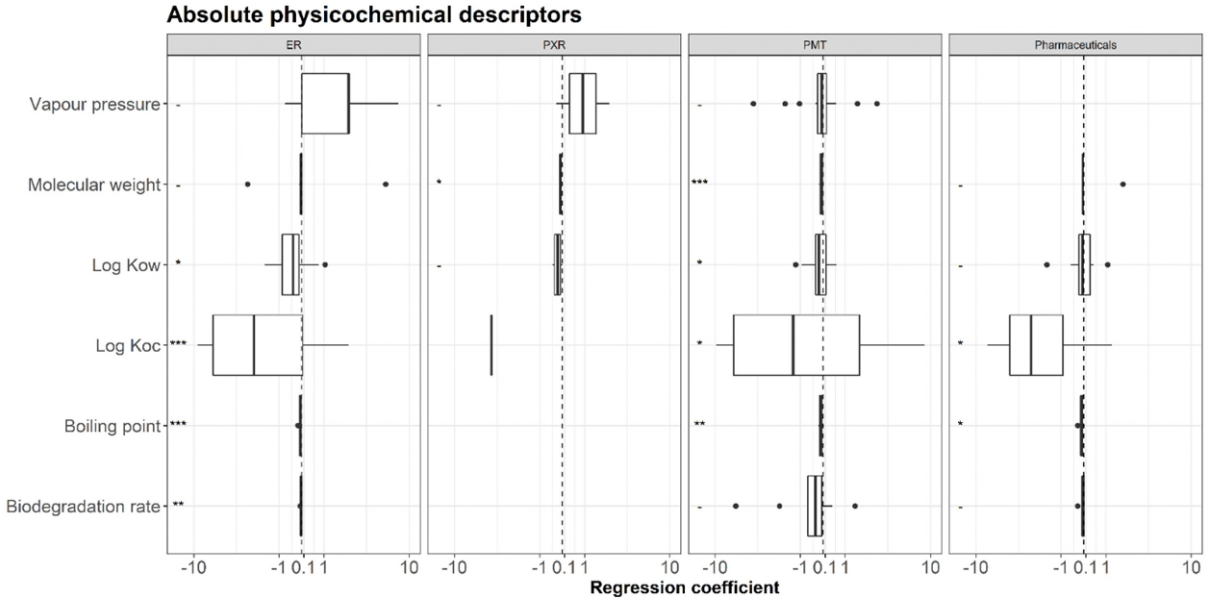


*Figure 5: Boxplots depicting the distribution of regression coefficients (minimum value, 25th percentile, median value, 75th percentile and maximum value) for vapor pressure (in mmHg), molecular weight (in g/mol), octanol-water partition coefficient (log Kow), sorption coefficient (log Koc), boiling point (in °C), and biodegradation rate (half-life in days) resulting from the random forest analysis for the four subsets of chemicals. ***; $\mu \neq 0$, $p < 0.001$, **; $\mu \neq 0$, $p < 0.01$, *; $\mu \neq 0$, $p < 0.05$, -; $\mu = 0$, $p > 0.05$*

2.4.4 Are PMOCs less toxic?

It was hypothesized that the properties that determine mobility make PMOCs inherently less toxic than non-polar compounds. A dataset was curated which contains chemical properties and *in vitro* toxicity data for 3360 chemicals relevant to the water cycle. Using random forest analyses and linear regressions, it was shown that the properties of compounds relating to mobility (K_{OC} and K_{OW}) are among the most prominent and often significant descriptors of toxicity and are inversely related to measured toxicity expressed as AC_{50}. This was also the case for three out of four substance groups. However, for PXR-active compounds, no significant correlations could be found between descriptors related to persistence and mobility, and toxicity. This might be due to the relatively small dataset and/or to the promiscuous nature of the PXR, as it binds a wide variety of ligands spanning a range of polarities. Ideally, for each individual toxicity test included, each physicochemical characteristic should consist of at least 30 data entries, after which the sampling distribution approximates the standard normal distribution (Kwak & Kim, 2017). Unfortunately, this was not the case for all *in vitro* tests in the subgroup with PXR-active compounds.

In the present study, the data were truncated based on a traditional cut-off point for log K_{ow} of 5 in environmental toxicology (Jonker & Van der Heijden, 2007). However, controlling the exposure in *in vitro* assays is challenging even for moderately hydrophobic due to their binding to the medium components and well plate walls in small test systems. A study by Jahnke et al. (2016) clearly illustrated that for chemicals beyond a log K_{OW} of 3, the freely-dissolved concentrations in the exposure medium can drop significantly relative to the nominal concentration. This could lead to the underestimation of the intrinsic toxicity of these chemicals, due to a higher nominal than actual concentration in the toxicity test. Such an underestimation could, in turn, obscure potentially observed relationships between chemical mobility and toxicity. Nonetheless, Jahnke et al. (2016) demonstrate, using a mass balance model by Armitage et al. (2014), that the cellular concentrations of chemicals remain quite stable over a wide range of K_{OW}, despite the substantial decrease in their freely-dissolved concentrations with increasing K_{OW}. Interestingly, despite any bias that may have arisen from the partitioning of more hydrophobic chemicals within the *in vitro* test systems used in ToxCast, as well as the exclusion of very hydrophobic compounds from the dataset in the set of PMT-compounds, both the random forest analyses and the linear regressions elucidated convincing relationships between mobility and toxicity in the present study.

Next to mobility, persistence also appeared to be a significant predictor of intrinsic toxicity for the total dataset. This significant correlation was, however, lost when of data with only few very biodegradable substances were tested (PMT compounds, pharmaceuticals). Therefore, it can be argued that this observation might not be related to the intrinsically higher toxicity (lower AC_{50}) of more persistent chemicals. Rather, this observation was likely related to the loss of readily biodegradable compounds during exposure in the toxicity tests, resulting in lower integrated exposure concentrations over time, leading to the overestimation of AC_{50} values when using nominal or initial concentrations as reference point (ECHA, 2017).

Hence, it is concluded that mobile organic compounds (MOCs) are indeed less toxic than similar, but more hydrophobic chemicals, owing to their mobility. This knowledge can aid in the prioritization of research and remediation efforts. Less persistent organic compounds are probably not intrinsically less toxic but pose a lower threat as the exposure duration and/or concentration is lowered by degradation and/or metabolism. This highlights that persistent and mobile chemicals, by definition, pose a threat because their persistence and mobility will, in due time, lead to accumulation in the (aqueous) environment and the (chronic) exposure of organisms, even far from original sources and hotspots. This makes their environmental emission, especially over extended periods, intrinsically problematic (Cousins et al., 2022; Scheringer et al., 2022).

Despite the presumed lower toxicity of PMOCs compared to less mobile compounds, the risk they could pose to humans and the environment should not be discounted. Chemical risk is a product of the hazard of a chemical and the exposure of organisms to that chemical (van Leeuwen & Vermeire, 2007). Although the hazards (i.e. intrinsic toxicities) of MOCs are presumably lower than those of their hydrophobic counterparts, their pervasiveness in water resources and drinking water results in continuous exposure in or through these matrices, which could contribute to

substantial risks (Hale et al., 2020). Rather, the findings presented here should highlight the relevance of P (persistence) and M (mobility) as compound properties in the hazard classification of chemicals, the interpretation and definition of exposure concentrations in bioassays, and the translation from *in vitro* to *in vivo* toxicity, and especially underline the data gap that exists regarding PMOC toxicities (Hale et al., 2022).

2.4.5 Closing knowledge gaps on PMOC toxicity

The presently obtained understanding of PMOC toxicity and its relation to chemical properties can inform frameworks for the identification of toxic PMOCs. After all, not all mobile chemicals are of lower toxicity, which is evident from the existence of PMT chemicals. Between PMOCs, these PMTs are expected to pose the most severe threat to humans and the environment. Therefore, gaining insight into what defines PMTs within the larger chemical universe of PMOCs can aid the effective regulation of these chemicals of concern. The elevated toxicity of certain chemicals compared to others is likely related to particular chemical structures or features that give rise to toxic activity via particular modes of action, of which two have been addressed in this study. Hence, an early warning system for PMTs that identifies presently overlooked or unknown chemicals that may pose particular risks can be based on the presence of such particular structures, features, or combinations of those. The present study, although providing insight into relationships between chemical properties and toxicity, did not yet provide sufficient information for such an early warning system. A scoring system for PMT identification was recently developed that incorporates a combination of qualitative and quantitative information on physicochemical properties and molecular structures of chemicals (Hartmann et al., 2022). This provides a valuable step towards the identification and classification of PMT chemicals in databases like REACH. However, a fully quantitative early warning model is yet to be developed, which can be based on the results presented here. The mechanistic understanding of PMOC toxicity that such models will provide, combined with measured toxicity data from high-throughput *in vitro* assays as were used in the present study, enables the closing of the knowledge gap on PMOC toxicity. Moreover, this will also provide much needed insight into additive or even synergistic effects of the mixtures of PMTs that are pervasive in present day water resources and cycles (Neuwald et al. 2022). Together, this can aid the prospective risk management of PMOCs which is needed to fulfil the European Union's ambition of a toxic-free environment by the year 2050, also for those chemicals that are most difficult to remove from aqueous matrices.

2.5 Acknowledgements

This work was performed under the Joint Research (BTO) programme of the Dutch water utilities, the Flemish utility De Watergroep, and Vewin. The authors gratefully acknowledge the funding received from the BTO programme. The authors thank Eric Verbruggen, Julia Hartmann and Emiel Rorije (RIVM), Rosa Sjerps (Oasen), and Mace Barron and Faith Lambert (US EPA) for their valuable input during the conception, analysis and writing of the manuscript.

3 References

Albergamo, V., Blankert, B., Cornelissen, E. R., Hofs, B., Knibbe, W.-J., van der Meer, W., & de Voogt, P. (2019). Removal of polar organic micropollutants by pilot-scale reverse osmosis drinking water treatment. *Water research, 148*, 535-545. doi:https://doi.org/10.1016/j.watres.2018.09.029

Armitage, J. M., Wania, F., & Arnot, J. A. (2014). Application of mass balance models and the chemical activity concept to facilitate the use of in vitro toxicity data for risk assessment. *Environmental science & technology, 48*(16), 9770-9779.

Arp, H. P. H., & Hale, S. E. (2022). Assessing the Persistence and Mobility of Organic Substances to Protect Freshwater Resources. *ACS Environmental Au*. doi:10.1021/acsenvironau.2c00024

Bajard, L., Adamovsky, O., Audouze, K., Baken, K., Barouki, R., Beltman, J. B., Beronius, A., Bonefeld-Jørgensen, E. C., Cano-Sancho, G., de Baat, M. L., Di Tillio, F., Fernández, M. F., FitzGerald, R. E., Gundacker, C., Hernández, A. F., Hilscherova, K., Karakitsios, S., Kuchovska, E., Long, M., Luijten, M., Majid, S., Marx-Stoelting, P., Mustieles, V., Negi, C. K., Sarigiannis, D., Scholz, S., Sovadinova, I., Stierum, R., Tanabe, S., Tollefsen, K. E., van den Brand, A. D., Vogs, C., Wielsøe, M., Wittwehr, C., & Blaha, L. (2023). Application of AOPs to assist regulatory assessment of chemical risks – Case studies, needs and recommendations. *Environmental Research, 217*, 114650. doi:https://doi.org/10.1016/j.envres.2022.114650

Bakand, S., & Hayes, A. (2010). Troubleshooting methods for toxicity testing of airborne chemicals in vitro. *Journal of Pharmacological and Toxicological Methods, 61*(2), 76-85. doi:https://doi.org/10.1016/j.vascn.2010.01.010

Becker, C. D., & Crass, D. W. (1982). Examination of procedures for acute toxicity tests with the fathead minnow and coal synfuel blends. *Archives of Environmental Contamination and Toxicology, 11*(1), 33-40. doi:10.1007/BF01055183

Bil, W., Govarts, E., Zeilmaker, M. J., Woutersen, M., Bessems, J., Ma, Y., Thomsen, C., Haug, L. S., Lignell, S., Gyllenhammar, I., Palkovicova Murinova, L., Fabelova, L., Tratnik, J. S., Kosjek, T., Gabriel, C., Sarigiannis, D., Pedraza-Diaz, S., Esteban-López, M., Castaño, A., Rambaud, L., Riou, M., Franken, C., Colles, A., Vogel, N., Kolossa-Gehring, M., Halldorsson, T. I., Uhl, M., Schoeters, G., Santonen, T., & Vinggaard, A. M. (2023). Approaches to mixture risk assessment of PFASs in the European population based on human hazard and biomonitoring data. *International Journal of Hygiene and Environmental Health, 247*, 114071. doi:https://doi.org/10.1016/j.ijheh.2022.114071

Breen, M., Ring, C. L., Kreutz, A., Goldsmith, M.-R., & Wambaugh, J. F. (2021). High-throughput PBTK models for in vitro to in vivo extrapolation. *Expert opinion on drug metabolism & toxicology, 17*(8), 903-921.

Brieuc, M. S., Waters, C. D., Drinan, D. P., & Naish, K. A. (2018). A practical introduction to Random Forest for genetic association studies in ecology and evolution. *Molecular ecology resources, 18*(4), 755-766.

Cousins, I. T., Johansson, J. H., Salter, M. E., Sha, B., & Scheringer, M. (2022). Outside the Safe Operating Space of a New Planetary Boundary for Per- and Polyfluoroalkyl Substances (PFAS). *Environmental science & technology, 56*(16), 11172-11179. doi:10.1021/acs.est.2c02765

Dix, D. J., Houck, K. A., Martin, M. T., Richard, A. M., Setzer, R. W., & Kavlock, R. J. (2007). The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicological sciences, 95*(1), 5-12.

Droge, S. T. (2018). Membrane–water partition coefficients to aid risk assessment of perfluoroalkyl anions and alkyl sulfates. *Environmental science & technology, 53*(2), 760-770.

ECHA. (2017). Guidance on information requirements and chemical safety assessment Chapter R. 7a: Endpoint specific.

Escher, B., & Schwarzenbach, R. (1995). Partitioning of substituted phenols in liposome– water, biomembrane– water, and octanol– water systems. *Environmental science & technology, 30*(1), 260-270.

Escher, S. E., Partosch, F., Konzok, S., Jennings, P., Luijten, M., Kienhuis, A., de Leeuw, V., Reuss, R., Lindemann, K.-M., & Bennekou, S. H. (2022). Development of a Roadmap for Action on New Approach Methodologies in Risk Assessment. *EFSA Supporting Publications, 19*(6), 7341E. doi:https://doi.org/10.2903/sp.efsa.2022.EN-7341

European Comission. (2020). *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS: Chemicals Strategy for Sustainability Towards a Toxic-Free Environment*. Retrieved from

European Commission. (2022a). Commission Review of REACH Annexes. Retrieved from https://ec.europa.eu/environment/chemicals/reach/reviews_en.htm#annex13

European Commission. (2022b). Hazardous chemicals – updated rules on classification, labelling and packaging. Retrieved from https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13578-Hazardous-chemicals-updated-rules-on-classification-labelling-and-packaging_en

Fischer, F. C., Henneberger, L., König, M., Bittermann, K., Linden, L., Goss, K.-U., & Escher, B. I. (2017). Modeling Exposure in the Tox21 in Vitro Bioassays. *Chemical research in toxicology, 30*(5), 1197-1208. doi:10.1021/acs.chemrestox.7b00023

Gollong, G., Neuwald, I. J., Kuckelkorn, J., Junek, R., & Zahn, D. (2022). Assessing the protection gap for mobile and persistent chemicals during advanced water treatment – A study in a drinking water production and wastewater treatment plant. *Water research, 221*, 118847. doi:https://doi.org/10.1016/j.watres.2022.118847

Groothuis, F. A., Heringa, M. B., Nicol, B., Hermens, J. L. M., Blaauboer, B. J., & Kramer, N. I. (2015). Dose metric considerations in in vitro assays to improve quantitative in vitro–in vivo dose extrapolations. *Toxicology, 332*, 30-40. doi:https://doi.org/10.1016/j.tox.2013.08.012

Hale, S. E., Arp, H. P. H., Schliebner, I., & Neumann, M. (2020). Persistent, mobile and toxic (PMT) and very persistent and very mobile (vPvM) substances pose an equivalent level of concern to persistent, bioaccumulative and toxic (PBT) and very persistent and very bioaccumulative (vPvB) substances under REACH. *Environmental Sciences Europe, 32*(1), 155. doi:10.1186/s12302-020-00440-4

Hale, S. E., Neumann, M., Schliebner, I., Schulze, J., Averbeck, F. S., Castell-Exner, C., Collard, M., Drmač, D., Hartmann, J., Hofman-Caris, R., Hollender, J., de Jonge, M., Kullick, T., Lennquist, A., Letzel, T., Nödler, K., Pawlowski, S., Reineke, N., Rorije, E., Scheurer, M., Sigmund, G., Timmer, H., Trier, X., Verbruggen, E., & Arp, H. P. H. (2022). Getting in control of persistent, mobile and toxic (PMT) and very persistent and very mobile (vPvM) substances to protect water resources: strategies from diverse perspectives. *Environmental Sciences Europe, 34*(1), 22. doi:10.1186/s12302-022-00604-4

Hammer, J. (2019). *Linking environmental properties of surfactants to molecular interactions.* Utrecht University.

Hartmann, J., Rorije, E., Wassenaar, P. N., & Verbruggen, E. (2022). Screening and prioritising Persistent, Mobile and Toxic chemicals: development and application of a robust scoring system.

Hermens, J. L., Heringa, M. B., & Ter Laak, T. L. (2007). Bioavailability in dose and exposure assessment of organic contaminants in (eco) toxicology. *Journal of Toxicology and Environmental Health, Part A, 70*(9), 727-730.

Hong, H., Tong, W., Fang, H., Shi, L., Xie, Q., Wu, J., Perkins, R., Walker, J. D., Branham, W., & Sheehan, D. M. (2002). Prediction of estrogen receptor binding for 58,000 chemicals using an integrated system of a tree-based model with structural alerts. *Environmental Health Perspectives, 110*(1), 29-36.

Hutchinson, T., Hellebust, J., Tam, D., Mackay, D., Mascarenhas, R., & Shiu, W. (1980). The correlation of the toxicity to algae of hydrocarbons and halogenated hydrocarbons with their physical-chemical properties *Hydrocarbons and halogenated hydrocarbons in the aquatic environment* (pp. 577-586): Springer.

Jahnke, A., Mayer, P., Schäfer, S., Witt, G., Haase, N., & Escher, B. I. (2016). Strategies for Transferring Mixtures of Organic Contaminants from Aquatic Environments into Bioassays. *Environmental science & technology, 50*(11), 5424-5431. doi:10.1021/acs.est.5b04687

Janitza, S., & Hornung, R. (2018). On the overestimation of random forest's out-of-bag error. *PloS one, 13*(8), e0201904.

Jonker, M. T., & Van der Heijden, S. A. (2007). Bioconcentration factor hydrophobicity cutoff: An artificial phenomenon reconstructed. *Environmental science & technology, 41*(21), 7363-7369.

Kavlock, R. J., Bahadori, T., Barton-Maclaren, T. S., Gwinn, M. R., Rasenberg, M., & Thomas, R. S. (2018). Accelerating the pace of chemical risk assessment. *Chemical research in toxicology, 31*(5), 287-290.

Krewski, D., Acosta Jr, D., Andersen, M., Anderson, H., Bailar III, J. C., Boekelheide, K., Brent, R., Charnley, G., Cheung, V. G., & Green Jr, S. (2010). Toxicity testing in the 21st century: a vision and a strategy. *Journal of Toxicology and Environmental Health, Part B, 13*(2-4), 51-138.

Kwak, S. G., & Kim, J. H. (2017). Central limit theorem: the cornerstone of modern statistics. *Korean journal of anesthesiology, 70*(2), 144-156.

Lambert, F. N., Vivian, D. N., Raimondo, S., Tebes-Stevens, C. T., & Barron, M. G. (2022). Relationships Between Aquatic Toxicity, Chemical Hydrophobicity, and Mode of Action: Log Kow Revisited. *Archives of Environmental Contamination and Toxicology*. doi:10.1007/s00244-022-00944-5

Mansouri, K., Grulke, C. M., Judson, R. S., & Williams, A. J. (2018). OPERA models for predicting physicochemical properties and environmental fate endpoints. *Journal of cheminformatics, 10*(1), 1-19.

Mitchell, M. W. (2011). Bias of the Random Forest out-of-bag (OOB) error for certain input parameters. *Open Journal of Statistics, 2011*.

Montes, R., Rodil, R., Cela, R., & Quintana, J. B. (2019). Determination of persistent and mobile organic contaminants (PMOCs) in water by mixed-mode liquid chromatography–tandem mass spectrometry. *Analytical chemistry, 91*(8), 5176-5183.

Næs, T., & Mevik, B. H. (2001). Understanding the collinearity problem in regression and discriminant analysis. *Journal of Chemometrics: A Journal of the Chemometrics Society, 15*(4), 413-426.

NCBI. (2022). PubChem.

Neumann, M., & Schliebner, I. (2019). *Protecting the sources of our drinking water: The criteria for identifying persistent, mobile and toxic (PMT) substances and very persistent and very mobile (vPvM) substances under EU Regulation REACH (EC) No 1907/2006*. Retrieved from

Neuwald, I. J., Hübner, D., Wiegand, H. L., Valkov, V., Borchers, U., Nödler, K., Scheurer, M., Hale, S. E., Arp, H. P. H., & Zahn, D. (2022). Occurrence, Distribution, and Environmental Behavior of Persistent, Mobile, and Toxic (PMT) and Very Persistent

and Very Mobile (vPvM) Substances in the Sources of German Drinking Water. *Environmental science & technology, 56*(15), 10857-10867. doi:10.1021/acs.est.2c03659

Proença, S., Escher, B. I., Fischer, F. C., Fisher, C., Grégoire, S., Hewitt, N. J., Nicol, B., Paini, A., & Kramer, N. I. (2021). Effective exposure of chemicals in in vitro cell systems: A review of chemical distribution models. *Toxicology in Vitro, 73*, 105133.

Reemtsma, T., Berger, U., Arp, H. P. H., Gallard, H., Knepper, T. P., Neumann, M., Quintana, J. B., & Voogt, P. d. (2016). Mind the Gap: Persistent and Mobile Organic Compounds—Water Contaminants That Slip Through. *Environmental science & technology, 50*(19), 10308-10315. doi:10.1021/acs.est.6b03338

Scheringer, M., Johansson, J. H., Salter, M. E., Sha, B., & Cousins, I. T. (2022). Stories of Global Chemical Pollution: Will We Ever Understand Environmental Persistence? *Environmental science & technology*. doi:10.1021/acs.est.2c06611

Schulze, S., Sättler, D., Neumann, M., Arp, H. P. H., Reemtsma, T., & Berger, U. (2018). Using REACH registration data to rank the environmental emission potential of persistent and mobile organic chemicals. *Science of the total environment, 625*, 1122-1128. doi:https://doi.org/10.1016/j.scitotenv.2017.12.305

Schulze, S., Zahn, D., Montes, R., Rodil, R., Quintana, J. B., Knepper, T. P., Reemtsma, T., & Berger, U. (2019). Occurrence of emerging persistent and mobile organic contaminants in European water samples. *Water research, 153*, 80-90. doi:https://doi.org/10.1016/j.watres.2019.01.008

Sjerps, R. M. A., Vughs, D., van Leerdam, J. A., ter Laak, T. L., & van Wezel, A. P. (2016). Data-driven prioritization of chemicals for various water types using suspect screening LC-HRMS. *Water research, 93*, 254-264. doi:https://doi.org/10.1016/j.watres.2016.02.034

U.S. EPA. (2012). *Estimating Physical / Chemical and Environmental Fate Properties with EPI Suite*. Retrieved from

U.S. EPA. (2015). *ToxCast & Tox21 Summary Files from invitrodb_v3*. Retrieved from: https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data

US EPA. (2022). EPI Suite™-Estimation Program Interface. Retrieved from https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface

van Leeuwen, C. J., & Vermeire, T. G. (2007). *Risk assessment of chemicals: an introduction* (Vol. 94): Springer.

Verhaar, H. J., Van Leeuwen, C. J., & Hermens, J. L. (1992). Classifying environmental pollutants. *Chemosphere, 25*(4), 471-491.

Watkins, R. E., Wisely, G. B., Moore, L. B., Collins, J. L., Lambert, M. H., Williams, S. P., Willson, T. M., Kliewer, S. A., & Redinbo, M. R. (2001). The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science, 292*(5525), 2329-2333.

Williams, A. J., Grulke, C. M., Edwards, J., McEachran, A. D., Mansouri, K., Baker, N. C., Patlewicz, G., Shah, I., Wambaugh, J. F., & Judson, R. S. (2017). The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. *Journal of cheminformatics, 9*(1), 1-27.

Wu, B., Li, S., & Dong, D. (2013). 3D structures and ligand specificities of nuclear xenobiotic receptors CAR, PXR and VDR. *Drug Discovery Today, 18*(11), 574-581. doi:https://doi.org/10.1016/j.drudis.2013.01.001

Yoon, M., Campbell, J. L., Andersen, M. E., & Clewell, H. J. (2012). Quantitative in vitro to in vivo extrapolation of cell-based toxicity assay results. *Critical reviews in toxicology, 42*(8), 633-652.

Zhou, J., Saeidi, N., Wick, L. Y., Kopinke, F.-D., & Georgi, A. (2021). Adsorption of polar and ionic organic compounds on activated carbon: Surface chemistry matters. *Science of the total environment, 794*, 148508. doi:https://doi.org/10.1016/j.scitotenv.2021.148508