Toxikologiska rådet
– expertorgan för rådgivning och samråd i toxikologiska frågor
The Toxicological Council
– body of experts for advice and consultation on toxicological issues

REPORT 1/22

Research report 2022

Methods for early identification of chemicals that have the potential to harm human health or the environment

Improvement of workflow and application for identification of NERCs in the Danish EPA database on articles

Preface

The Toxicological Council is an expert organisation established to facilitate the rapid identification of chemical substances that can be harmful to human health or the environment. The Council includes representatives from governmental authorities and academic institutions. The Toxicological Council identifies and evaluates signals of new, potential and emerging chemical risks and reports its findings to SamTox. This project was conducted as a consultancy commission in order to develop methodologies for identification of new or emerging chemical risks.

The report was written by Suzanne Bruks, Ziye Zheng and Prof. Patrik Andersson from the Department of Chemistry at Umeå University.

The project reference group consisted of Lina Wendt-Rasch and Olof Johansson, from the Swedish Chemicals Agency (KemI). Lina Wendt-Rasch was the project leader and contact person at KemI, which financed the project.

The Swedish Chemicals Agency financed the project in order to support the Toxicological council's assignment to provide updated and relevant information to SamTox. The conclusions presented in the report represent the views of authors and do not necessarily reflect the opinions of individual authorities and academic institutions in the Toxicological council.

In case you experience accessibility problems (i.e., tables), please contact KemI for assistance, either by e-mail to kemi@kemi.se or by calling us on +46 8 519 411 00.

Content

Sammanfattning

Under 2020 påbörjades ett projekt i syfte att inleda utveckling av metodik som möjliggör att utifrån stora datamängder, såsom det svenska produktregistret (SE-PR) eller storskaliga kartläggningar av kemikalier i olika material, identifiera och prioritera ämnen med potentiellt hälso- och miljöfarliga egenskaper. Projektet finansierades av Kemikalieinspektionen och utfördes av forskare vid Umeå universitet samt Sveriges lantbruksuniversitet (SLU). Inom ramen för projektet togs en metodik fram för datafiltrering och databehandling (datakurering) följt av egenskapsbaserad prioritering samt litteratur- och databasgranskning för att undersöka förekomst i framför allt miljöprover. De modellerade egenskaperna omfattade persistens (P), bioackumulation (B) och mobilitet (M) samt till viss del toxicitet (T). I den nuvarande studien har en vidareutveckling och utvidgning av detta projekt genomförts. Projektet finansierades av Kemikalieinspektionen och utfördes av forskare vid Umeå universitet.

Den arbetsmodell för semiautomatisk datakurering som togs fram under 2020 vidareutvecklades och applicerades på kemikalieinventeringar i vardagsprodukter som gjorts av den danska Miljöstyrelsen. Totalt ingick drygt 9000 dataposter i analysen från studier inriktade på kemikalier i produkter med hög potentiell exponering för konsumenter, eller produkter som misstänks innehålla farliga kemikalier. De danska studierna omfattar både ”target screening” där specifika ämnen analyserades i en mängd olika produkter och ”non-target” kartläggning av t.ex. kemiska grupper inom specifika produkttyper. Exempel på produkter som analyserades är 3D-printermaterial, cykelhjälmar, barnleksaker och kosmetika. De framtagna ingångssubstanserna bearbetades i datakureringsmodellen, vilket slutligen genererade 1 797 unika CAS-nummer och som i sin tur gav 1 361 unika kemiska strukturer lämpliga för miljö- och hälsobedömning via struktur-aktivitets modellering (QSAR). I det sista steget förkastades främst felaktiga strukturer, oorganiska ämnen, polymerer eller ämnen med element som är olämpliga för modellering.

Kurerade data från den danska databasen applicerades i QSAR-modeller för att prediktera persistens (P; resistens mot nedbrytning), bioackumulation (B; upptag i biota), mobilitet (M; rörlighet i vattenmiljön) och ett stort antal toxicitets-parametrar (T). Gränsvärden för persistens och bioackumulation influerade av Reach användes för att identifiera och prioritera potentiellt farliga ämnen. Filtrering av kemikalierna gjordes även med ansatta värden för kemikaliernas mobilitet och toxicitet. För att ta hänsyn till osäkerheter i modellerade värden jämfört med experimentellt framtagna värden användes även ett konservativt angreppssätt med lägre gränsvärden.

Projektet har totalt sett genererat kurerade strukturer och data som kan filtreras utifrån användar- och projektspecifika krav. Totalt identifierades 18 ämnen som vPvB (konservativ filtrering av P och B) och 24 ämnen som vPvM (konservativ filtrering av P och M) och dessa domineras av högfluorerade kemikalier (PFAS) men innehåller även t.ex. ett antal pigment.

Genererade listor med potentiella NERCs behöver studeras i ett nästa steg i detalj av expertis och kompletteras med data från experimentella studier och monitorering. Generellt sett är förekomstundersökningar baserade på litteratur och databaser ett krävande manuellt steg som behöver utvecklas för att ytterligare systematisera och effektivisera metoden.

1. Introduction

The Swedish Toxicological Council aims at early identification of chemicals of concern through expert-based identification and judgements. As a complement the Council has identified a need to develop a systematic and automatic tool, i.e., a methodology for and early warning systems of detection of New and Emerging Risk Chemicals (NERCs). This system should include a generic model for data curation and hazard screening and to be applicable to various datasets and chemical inventories. A pilot study with the objective to initiate such a system was completed 2020 funded by the Swedish Chemicals Agency.[1]

In the previous project it was concluded that the data curation, based on data mining from various open sources, is a crucial step to make inhomogeneous lists of chemical entries suitable for the hazard screening. The batch generation of reliable molecular structures in the form of simplified molecular input line-entry system (SMILES) generated by the data curation method makes it possible to produce large number of predictions from the hazard estimation models. However, these estimation models have limitations and a combined result from several models were used to increase the reliability.[2] Combining several estimation models provides two advantages; less susceptibility to limitations of any single model and consensus within the result.

The overall aim of this project was to further refine the methodology developed during 2020 for early detection of NERCs and apply it to the Danish EPAs dataset of everyday consumer products. This included data curation of the database and hazard scoring and ranking based on model estimates of measures of persistence, bioaccumulation potential, mobility and toxic properties. We also aimed at improving generated lists of NERCs from the previous project and complement those lists with toxicity hazard scores. Focus in the project was the data curation and scoring based on PB and M whereas the T scoring warrants further investigations.

2. Material and methods

2.1 Databases and inventories

The Ministry of Environment of Denmark (Danish EPA) has constructed a database of chemicals found in consumer products analysed in surveys since 2001.[3] At the time for this report the database consists of more than 9000 entries from 186 surveys. The surveys target substances in a large range of products with high exposure to humans, or products that are suspected to contain hazardous chemicals. Hence, the surveys contain both target screening where specific compounds are investigated in a variety of products, and mapping of specific chemical groups within specific product types. The database was curated in 2020, however four additional surveys have been published during 2021 and therefore the data curation was repeated for this project. In addition, the curated dataset was used for hazard screening.

In addition, two databases were revisited and more information on the toxicity of the included chemicals were added, in line with data from the databases from Danish EPA. These two are 1) an inventory of additives in plastics and 2) chemicals used in the paper and paperboard industry. These are both described in more detail in the previous report.[1]

2.2 Data curation

The general methodology for data curation and hazard screening as previously described was applied for this project with minor revisions as described below (Figure 1).[1] The latest version of the open-source software Konstanz Information Miner[4] (KNIME (4.4.1)) was downloaded and updated with all available extensions, including Palladian Nodes for KNIME[5] accessed online. The publicly available data curation workflow used for this project[6] (version July 2021) was imported to the KNIME analytics platform. The updates include assorting polymers based on the abbreviation “poly”, as suggested for improvement by the previous project.[1,6]

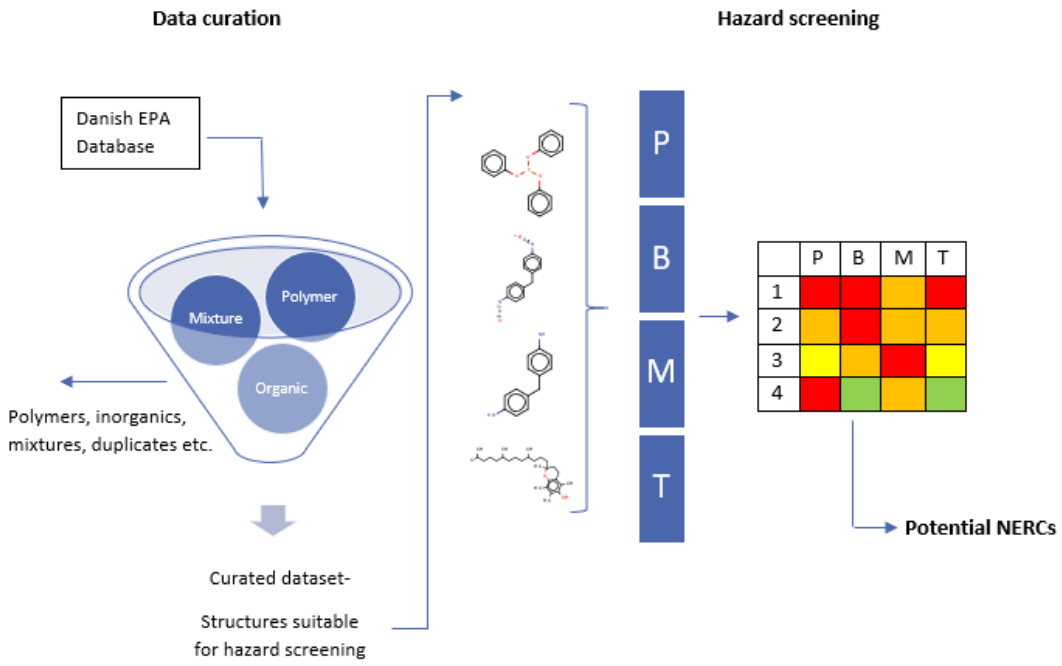

Figure 1. Overview of the methodology applied in the project.

In accordance with the previous project,[1] the dataset from the Danish EPA was manually reviewed, duplicates removed and CAS numbers corrected, in preparation for data curation. Five positive controls (Table 1) with well characterized PB properties[7–11] were added to the dataset before hazard screening.

Table 1. Positive controls added to the dataset with well characterized hazardous properties

Name	CAS	SMILES
HBCDD	3194-55-6	BrC1CCC(Br)C(Br)CCC(C(CCC1Br)Br)Br
Triclosan	3380-34-5	Clc1ccc(c(c1)O)Oc1ccc(cc1Cl)Cl
DDT	50-29-3	ClC(C(c1ccc(cc1)Cl)c1ccc(cc1)Cl)(Cl)Cl
TBBPA	79-94-7	CC(c1cc(Br)c(c(c1)Br)O)(c1cc(Br)c(c(c1)Br)O)C
Benzo[a]pyrene	50-32-8	c1ccc2c(c1)c1ccc3c4c1c(c2)ccc4ccc3

2.3 Hazard screening

2.3.1 Fate screening

For the hazard screening quantitative structure–activity relationships (QSARs) models were applied using the platforms EPI Suite[12] and VEGA[13] to predict persistence (P), bioconcentration (B), and mobility (M). Three models were used for bioconcentration and two models for persistence. To access mobility properties three models were used to predict the organic carbon-water partition coefficient (K_{oc}) and two models for water solubility, in accordance with German EPA (UBA)[14]. For qualification of M-properties at any hazard level the water solubility estimation has to be above 0.15 mg/L.

To allow for the construction of prioritization lists of the substances the QSAR results were scored using priority marks. The thresholds were influenced by REACH limits where "alarming" corresponds to very P, very B and very M[1], the "probable" level corresponds to P, B and M limits. The "possible" level was set to reach a conservative filter of possible P, B and M chemicals which are thresholds lower than the REACH limits (see Table 2). In addition, a numeric score was derived using the system as showed in Table 2 from 0 to 3 with increasing potential hazard.

Table 2. Threshold values for scoring chemicals with priority mark. Estimations of probable and alarming are above REACH qualification level of each endpoint.

Priority mark	**Persistence (P)**	**Bioaccumulation**	**Mobility**		**Scoring**
	Water (half-lives, days)	BCF (L/kg)	Log K_{oc} + S_w (mg/L)		
Alarming (red)	160	> 5000	< 3	≥ 0.15	3
Probable (orange)	60	2000 - 5000	>3- < 4	≥ 0.15	2
Possible (yellow)	32.5	500 - 2000	> 4 - < 6	≥ 0.15	1
Unlikely (green)	< 32.5	< 500	> 6		0

To weigh the different model outcomes for prioritization, two approaches were used. Either the "highest individual score" was used, where at least one model has indicated high hazard, or the "total score" was calculated allowing for prioritizing those substances where several models have indicated high hazards, see Table 3.

Table 3. Example of heat map scoring for hazard properties (exemplified for endpoints with three different models).

	X model 1	X model 2	X model 3	Total score (0-9)	Highest individual score
Compound 1	3	2	3	**8**	3
Compound 2	3	2	1	**6**	3
Compound 3	2	2	2	**6**	2
Compound 4	0	0	3	**3**	3

2.3.2 Toxicological screening

As a pilot case study human health and environmental hazard estimates related to toxic potential were calculated using the VEGA QSAR platform (version 1.1.5-b48) for the Danish database of chemicals found in consumer products as well as for the list of substances found in paper and plastic additives.[1] The platform includes models from several sources including CEASAR and TEST by USEPA. Note that these data are not discussed below but presented in the supplementary material.

Data from 34 models were used representing acute and chronic toxicity, skin sensitization, hepatotoxicity, endocrine disruption, or cancerogenic, mutagenic potential and developmental/reproductive toxicity (CMR). The reliability of the results from VEGA are categorized as; good reliability, moderate reliability, low reliability or experimental values (for information on the reliability categorization, see vegahub.eu).

To allow for the construction of prioritization lists of substances, threshold values were constructed to reach chemicals of "probable" concern and of "unlikely" concern. The settings of these thresholds were inspired by thresholds given in REACH and used by GreenScreen. The scoring was done differently for quantitative and qualitative outcomes of the QSAR models, for more information see the supplementary information.

2.4 Description of the Supplementary material

The database including chemical names, CAS numbers, SMILES, predicted hazard estimates reflecting PBT and M estimates and hazard scores are given in the supplementary information. The database enables sorting chemicals based on consensus score, individual score or hazard endpoints. Scoring is carried out for each individual CAS, however those compounds that have the same SMILES, share ID number in the NERC list, are marked as duplicates.

More details on the scoring system are presented in the Excel files entitled Supplementary information – Danish EPA database and Supplementary information – Paper Plastic.

3. Results

The data base of the Danish EPA included 9,499 entries that were condensed to a total of 1,797 individual CAS numbers (Figure 2). After data curation, 1,361 SMILES were identified suitable for hazard screening. The total structural yield was 81% after data curation. The batch of SMILES strings were used for estimation of hazard properties, given in the supplementary material, and used for screening and prioritizing potential NERCs. In addition to P, B and M related properties a range of T data was generated. These can be found in the supplementary file including notification on reliability and various means of condensing and scoring approaches. The T data is however not discussed here nor used in the filtering process aiming for identification of NERCs.

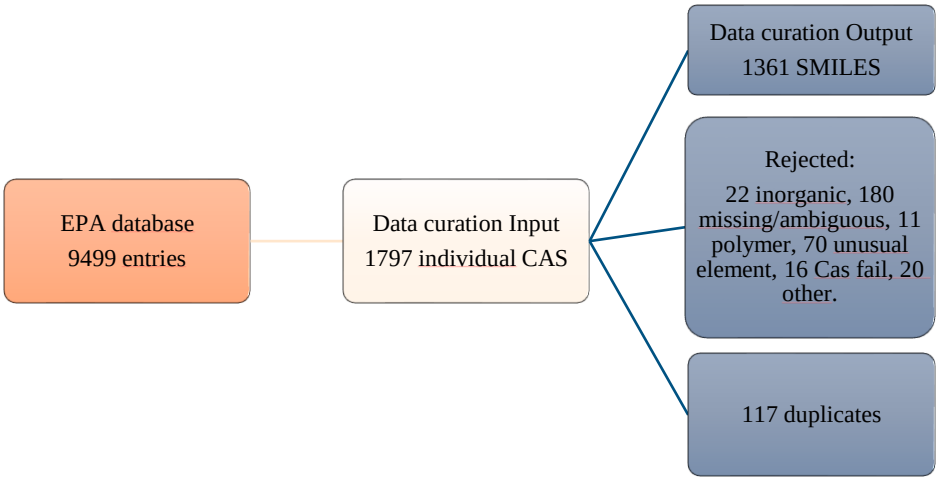

Figure 2. Quantitative result of the data curation of the Danish EPA database.

3.1 Fate screening

With a focus on P, B and M, any compound with a minimum score of 1 (i.e., at least one of the models resulted in the hazard mark "possible") for both P and B (n= 267) or P and M (n=551) are listed as possible NERCs (Figure 3). Note that substances categorized as "probable" and "alarming" are also included in the "possible" category". The number of potential PM and PB substances decreases substantially within the REACH limits, from 554 to 120 and from 267 to 43 substances, respectively.

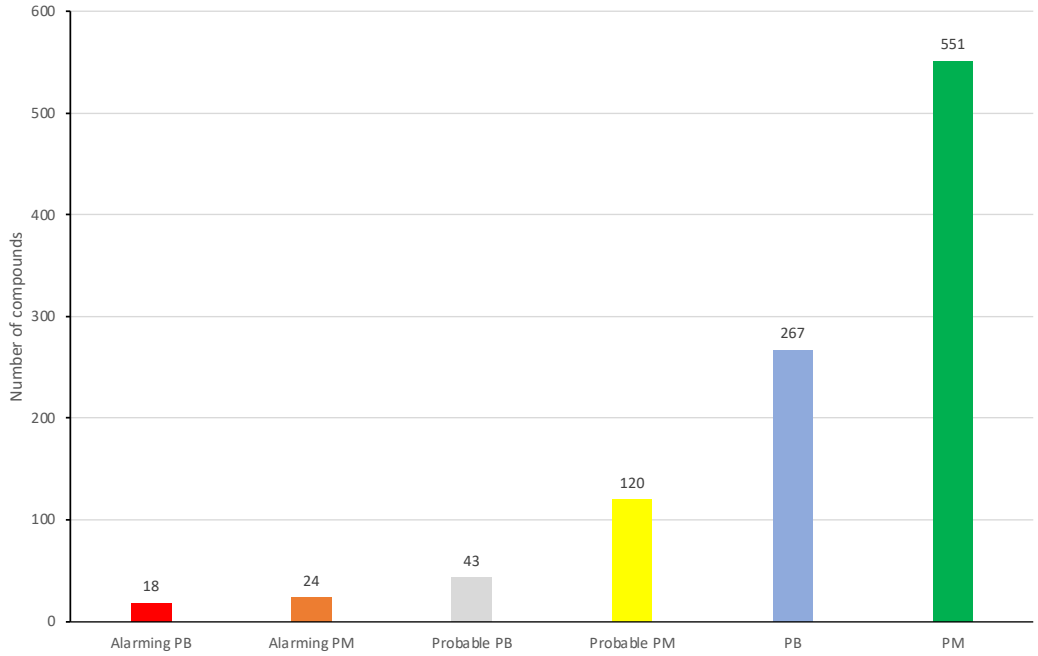

Figure 3. Number of compounds found for each hazard mark.

The substances with a hazard mark alarming for PB (n=18) and PM (n=24) in the hazard screening is presented in Table 5. Four of the substances are estimated both as alarming for both PB and PM. The name and year of the Danish survey(s) in

which the compounds have been detected are also given the Table 5.[15] Per- and polyfluoroalkyl substances (PFAS) are overrepresented with 83% of the alarming PB and 54% of alarming PM substances.

Table 5. Substances ranked as vPvB and vPvM by at least one of the hazard screening models for each endpoint, and their occurrence in surveys performed by the Danish EPA.

CAS	Name	Endpoints	Surveys
8054-98-6	Hexachlorophen	vPvB	Antibacterial agents in garments (2003)
481071-78-7	8:2 Fluorotelomer sulfonat	vPvB	Health assessment of fluorine substances in cosmetic products (2018) Identification and risk assessment of chemical substances in bicycle helmets (2018)
335-67-1	Perfluoroctansyra	vPvB vPvM	Identification and health assessment of chemical substances in shoe care products (2005) Exposure of chemical substances in waterproofing (2004) Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016) Health assessment of fluorine substances in cosmetic products (2018)
375-85-9	Perfluorheptansyra	vPvB vPvM	Identification and health assessment of chemical substances in shoe care products (2005) Exposure of chemical substances in waterproofing (2004) Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016) Health assessment of fluorine substances in cosmetic products (2018)
647-42-7	6:2 Fluortelomeralkohol	vPvB	Identification and health assessment of possible harmful components of textile impregnation sprays (2008) Identification, marketing and health assessment of chemical substances in baby products (2008) Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)
172155-07-6	Perfluoro(3,7-dimethyloctanoic acid)	vPvB	Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)
3864-99-1	ls-104191	vPvB	Identification, emissions and environmental and health assessment of chemical substances in artificial grass (2008)

CAS	Name	Endpoints	Surveys
79-94-7	TBBPA	vPvB	Identification, placing and health assessment of chemical substances in foam plastic toys and children's articles (2006) Chemical substances in carpets (2002)
1996-88-9	8:2fluortelomermethacrylat	vPvB	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015)
70887-84-2	8:2 Fluortelomercarboxylsyra	vPvB	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015)
70887-88-6	6:2 Fluortelomercarboxylsyra	vPvB vPvM	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015)
27905-45-9	8:2fluortelomeracrylat	vPvB	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018)
4151-50-2	N-Ethyl-perfluoroctane sulfonamide	vPvB	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)
31506-32-8	n-methylperfluoroctansulfonamid	vPvB	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)
1763-23-1	Perfluoroctansulfonsyra	vPvB vPvM	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)
24448-09-7	n-methylperfluorctan sulfonamidethanol	vPvB	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)
865-86-1	10:2fluortelomeralkohol	vPvB	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)

CAS	Name	Endpoints	Surveys
375-95-1	Perfluornonansyra	vPvB	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016) Health assessment of fluorine substances in cosmetic products (2018)
2706-28-7	CI 13015	vPvM	Dyes in tattoo marks (2005)
12225-21-7	CI 19140	vPvM	Mapping of chemical substances in halloween and theatrical make-up (2002)
80443-41-0	Tebuconazol	vPvM	Release of chemical substances from exotic wood products (2005)
6358-31-2	Pigment Yellow 74	vPvM	Study of dyes in tattoo colours (2002)
108-78-1	1,3,5-triazine-2,4,6-triamine, Melamin	vPvM	Exposure of 2-year-olds to chemical substances (2009) Chemical substances in child car seats and other products (2015)
414911-30-1	4:2 Fluorotelomer sulfonat	vPvM	Health assessment of fluorine substances in cosmetic products (2018) Identification and risk assessment of chemical substances in bicycle helmets (2018)
1546-95-8	7H-Dodecafluoroheptanoic acid	vPvM	Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016) Health assessment of fluorine substances in cosmetic products (2018)
425670-75-3	6:2 Fluorotelomer sulfonat	vPvM	Health assessment of fluorine substances in cosmetic products (2018) Identification and risk assessment of chemical substances in bicycle helmets (2018)
19363-60-1	5-Chlor-3-phenyl-4-isothiazolcarbonitril	vPvM	Risk assessment of 3D printers and 3D printed products (2017)
3089-11-0	Hexa(methoxymethyl)melamin	vPvM	Screening of health effects from chemical substances in textile colours (2005)
375-73-5	Perfluorbutansulfonsyra	vPvM	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)
2706-91-4	Perfluorpentansulfonsyra	vPvM	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015)
2043-47-2	4:2fluortelomeralkohol	vPvM	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)
12236-62-3	Pigmentorange36	vPvM	Study of dyes in tattoo colours (2002)

CAS	Name	Endpoints	Surveys
18472-51-0	Chlorhexidine Digluconate	vPvM	Mapping and health and environmental assessment of hand soap (2006)
307-24-4	Perfluorhexansyra	vPvM	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016) Health assessment of fluorine substances in cosmetic products (2018)
355-46-4	Perfluorhexansulfonsyra	vPvM	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification and risk assessment of chemical substances in carpets for children (2016)
78-43-3	Tris-dichlorpropylphosphat	vPvM	Chemical substances in child car seats and other products (2015)
13674-87-8	Tris(1,3-dichlor-2-propyl)phosphat	vPvM	Chemical substances in child car seats and other products (2015) Identification and risk assessment of chemical substances in bicycle helmets (2018) Identification, health and environmental assessment of flame retardants in textiles (2014)
52591-27-2	4:2fluortelomeracrylat	vPvM	Polyfluoroalkyl compounds (PFAS) in textiles for children (2015)

3.2 Toxicological screening

The results section is focused on the P, B and M hazard scoring, however, T was assessed and can be used for further ranking and identification of NERCs. T can be studied and scored depending on interest using individual scores of particular hazard outputs or clustered in e.g. endocrine disruption or CMR, or given as a summed T-scores, see supplementary material. T scores are provided with reliability scores which could be used for assessing quality of assessment.

4. Discussion and conclusion

One important factor for the result of QSAR estimations is the quality of the molecular structures, and hence a proper data curation method. Often a large percentage of the initial data is rejected in the data curation step. However, curation of the Danish list of chemicals in consumer products resulted in a large percentage (81%) of entries being kept. Likely this high yield in the curation, based on CAS numbers and names, is due to the nature of the dataset consisting of many chemical target analysis surveys with common compounds giving a dataset with few complex entries, such as mixtures and polymers. The dataset is expected to contain many risk chemicals since many of the surveys are targeting known problematic compounds. However, some of the surveys are rather old and may give a misleading image over which compounds are in current use. Notably, many of the alarming risk chemicals identified are PFAS chemicals, a large group of compounds that currently are under investigation.

In summary, this project has generated lists of potential NERCs from the Danish EPA database based on individual output data from a range of hazard estimation tools. In addition, a large range of measures of toxicity potential is given both for the Danish EPA database but also for chemicals on the previously curated databases of plastic additives and paper and paperboard chemicals.[1] These data warrants detailed analysis including ranking using reliability scores and consideration of applicability domains of individual models. We have developed the supplementary material with new means for sorting the chemicals including the highest individual and mutual score. This enables tailor made sorting of large number of chemicals depending on research focus of future studies.

5. References

1. Research report 2021. Methods for early identification of chemicals that have the potential to harm human health or the environment. https://www.kemi.se/download/18.719852d17dbbc5435a23f/1640165610485/Toxicological-council-Research-report-2021-Methods-for-early-identification-of-%20new-or-emerging-risk-chemicals.pdf

2. Gissi, A. *et al.* Integration of QSAR models for bioconcentration suitable for REACH. *The Science of the total environment* **456–457**, 325–332 (2013).

3. Danish surveys on consumer products. *Accessed: 2021-09-06* https://eng.mst.dk/chemicals/chemicals-in-products/consumers-consumer-products/danish-surveys-on-consumer-products/.

4. KNIME Analytics Platform | KNIME. https://www.knime.com/knime-analytics-platform.

5. Palladian — NodePit. *Downloaded September 2021* https://nodepit.com/product/palladian.

6. Gadaleta, D., Lombardo, A., Toma, C. & Benfenati, E. Correction to: A new semi-automated workflow for chemical data retrieval and quality checking for modeling applications. *Journal of cheminformatics* **11**, (2019).

7. Update CHANGELOG.txt · DGadaleta88/data_curation_workflow@dfc7697 · GitHub. *Read September 2021* https://github.com/DGadaleta88/data_curation_workflow/commit/dfc7697be778631ff29cd14735cc62c3165771e3.

8. Substance Information - ECHA. *Accessed 2021-11-11* https://echa.europa.eu/substance-information/-/substanceinfo/100.000.026.

9. Substance Information - ECHA. *Accessed 2021-11-11* https://echa.europa.eu/substance-information/-/substanceinfo/100.001.125.

10. Substance Information - ECHA. *Accessed 2021-11-11* https://echa.europa.eu/substance-information/-/substanceinfo/100.000.023.

11. Substance Information - ECHA. *Accessed 2021-11-11* https://echa.europa.eu/substance-information/-/substanceinfo/100.019.724.

12. Substance Information - ECHA. *Accessed 2021-11-11* https://www.echa.europa.eu/substance-information/-/substanceinfo/100.020.167.

13. EPI Suite™-Estimation Program Interface | US EPA. *Downloaded October 2021* https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface.

14. Benfenati, E., Manganaro, A. & Gini, G. VEGA-QSAR: AI inside a platform for predictive toxicology. 21–28 (2013) doi:10.2/JQUERY.MIN.JS.

15. Crookes, M. J. & Fisk, P. Evaluation of using mobility of chemicals in the environment to fulfil bioaccumulation criteria of the Stockholm Convention FINAL REPORT. *Company Number* 5758319, (2018).

16. Database of chemicals in consumer products. https://eng.mst.dk/chemicals/chemicals-in-products/consumers-consumer-products/database-of-chemicals-in-consumer-products/.

Toxikologiska rådet

– expertorgan för rådgivning och samråd i toxikologiska frågor

The Toxicological Council

– body of experts for advice and consultation on toxicological issues