

ECETOC TRAv3: An In-depth Comparison of Publicly Available Measurement Data Sets With Modelled Estimates of Occupational Inhalation Exposure to Chemicals

Nenad Savic¹, Jan Urbanus^{2,*}, Oliver Henschel³, Qiang Li⁴, Dave Marsh⁵, Chris Money⁶, Dook Noij⁷, Paul van de Sandt⁸, Joost van Rooij⁹ and Matthias Wormuth¹⁰

¹Center for Primary Care and Public Health (Unisanté), University of Lausanne, Route de la Corniche 2, 1066 Epalinges, Lausanne, Switzerland;

²Shell Product Stewardship, Belgian Shell N.V., Brussels, B-1000, Belgium;

³Corporate Health Management, BASF SE, Ludwigshafen am Rhein, 67056, Germany;

⁴Clariant Produkte (Deutschland) GmbH, Sulzbach am Taunus, 65843, Germany;

⁵ExxonMobil Biomedical Sciences Inc, ExxonMobil, Leatherhead KT22 8UX, UK;

⁶Cynara Consulting, Brockenhurst S042 7RX, UK;

⁷In Personal Capacity, Formerly Dow Global Industrial Hygiene Expertise Centre, 4531 EB Terneuzen, The Netherlands;

⁸In Personal Capacity; Formerly Shell Health Risk Science Team, Shell International B.V., The Hague, 2596HR, The Netherlands;

⁹Caesar Consult, 6503 CB, Nijmegen, The Netherlands;

¹⁰Syngenta Crop Protection AG, 4058 Basel, Switzerland

*Author to whom correspondence should be addressed. Tel.: +32-497-515446; e-mail: jan.urbanus@shell.com

Abstract

In this study, 129 exposure situations (ESs) with six or more measured inhalation exposures to dust from solids or vapour from liquids in occupational settings were compared with modelled European Centre of Ecotoxicology and Toxicology of Chemicals (ECETOC) targeted risk assessment tool, version 3 (TRAv3) estimates. The measurement data were extracted from previously published studies examining TRAv3 performance and pooled into a curated database. The comparison exercise focussed on the vapour exposure scenarios, as there were too few dust scenarios for a meaningful analysis of any required model corrections. A group of experts in the exposure modelling field retrieved and reviewed the input parameters used in these ESs. Where considered appropriate, modifications were applied to better match the input parameter definitions and the scope of applicability of the TRAv3. Differences and mean absolute error (MAE) were calculated between the log-transformed modelled exposure value and the 75th percentile of each measured data set and regression analysis was performed. The results indicated that the TRAv3 overestimated 80% of the measured data sets. Both over- and underestimations were mostly by factors 1–5. The calculated MAE for liquids was 0.7, indicating that on average the difference between the 75th percentile and the TRAv3 estimate was less than one order of magnitude. A multiple linear regression showed that some input parameters such as medium volatility, certain process categories (PROC), industrial setting, and the presence of local exhaust ventilation are associated with underestimations. The results of the regression analysis can be used by TRAv3 users to review the degree of over- or underestimation in their current exposure assessments, compared to the curated database. Although multiple linear regression is an appropriate methodology to characterize the TRAv3's performance, more data sets are still needed in view of some remaining data gaps. Nevertheless, the results of the current analysis are being used by ECETOC to further develop the tool as a suitably conservative screening tool for use in REACH chemical safety assessment of occupational exposure to chemicals.

Keywords: EU chemical regulation; exposure model; regulatory risk assessment; tiered exposure assessment

Received: April 21, 2022. Accepted: January 3, 2023.

© The Author(s) 2023. Published by Oxford University Press on behalf of the British Occupational Hygiene Society.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

What's Important About This Paper?

Exposure models are widely used in chemical registration in the European Union, so it is important to understand model performance to ensure appropriate worker protection. This study compared measured vapour exposures with those modelled by the ECETOC Targeted Risk Assessment Tool version 3. The model overestimated exposures for 80% of the measured exposure datasets, though modelled values were generally within a factor of 1–5 of measured values.

Introduction

Several occupational exposure models have been developed for chemical exposure assessment within the context of the Registration, Evaluation, Authorization, and restriction of CHemicals (REACH) legislation (ECHA, 2016). The European Centre of Ecotoxicology and Toxicology of Chemicals (ECETOC) released the first version (ECETOC 2004) of its target risk assessment (TRA) tool in 2004, which was later updated to the second (ECETOC 2009) (TRAv2; in 2009) and a third version (ECETOC 2012) (TRAv3; in 2012). With a few other models, such as EMKG-EXPO-TOOL (Kindler and Winteler 2010), MEASE and Stoffenmanager (Marquart *et al.*, 2008; Schinkel *et al.*, 2010), the European Chemicals Agency (ECHA) recommends TRAv3 to be used as a first-tier model (ECHA, 2016). These models are required to be sufficiently conservative—to overestimate the actual (i.e. measured) exposure at the workplace for the chemical safety assessment to err on the side of caution. The more complex, tier-2, approach is recommended when the risk is not demonstrated to be adequately controlled under the conditions of use by the first-tier modelling. Although since recently ECHA defines Stoffenmanager as a higher, 1.5, tier model, the only tier-2 model for predicting worker inhalation exposure is the Advanced REACH Tool (Schinkel *et al.*, 2011, Fransman 2013). Unlike the lower-tier models, higher-tier models require more contextual information on the exposure situation to be coded into its input parameters.

The validity of modelled exposure estimates is often questioned, and many studies have compared the models' predictions with measurements (Kupczewska-Dobacka *et al.*, 2011; Mc Donnell *et al.*, 2011; Ishii *et al.*, 2017; Savic *et al.*, 2017, 2019; van Tongeren *et al.*, 2017). The most comprehensive study, known as the ETEAM project (van Tongeren *et al.*, 2017), investigated the performance of the first-tier models using around 2000 inhalation exposure measurements. The study found 28% and 35% of the individual measurements for solids and liquids, respectively, were underestimated by the TRA's estimate, which represents the 75th percentile (P75) of the exposure distribution (Money *et al.*, 2014). Lee *et al.* (2019b) performed a similar study using around 300 individual exposure measurements collected at 18 workplaces in the USA. The study

found that 38% of the measurements of exposure to volatile liquids were underestimated by TRAv3. Due to insufficient data size, results were not reported for solids. Moreover, both studies investigated the degree of underestimation per input parameter to determine which input entries caused these underestimations. Lee *et al.* (2019b) found that the TRA underestimated more than 60% of the measurements for process categories (PROCs) 3, 10, and 15; the ETEAM reported PROC 7 and 14 as the least conservative PROCs, that is, underestimated more than 70% cases, while PROCs 5 and 19, although less underestimated, were still considered as insufficiently conservative. In addition, both studies reported the estimates for low volatility bands and professional settings to be more conservative than for industrial settings or higher volatility bands. The two studies, however, disagreed regarding the effect of local exhaust ventilation (LEV) when present. While ETEAM found that 74% of the exposure situations (ESs) with LEV were underestimated, the ESs with LEV was more conservative in the study conducted by Lee *et al.* (2019b). These projects, along with several of the other reported projects, compared individual measurement results with a TRA prediction and then examined the overall over- and under-estimations, stratified by the various TRA input parameters such as PROC and setting. However, TRA predictions are based on the upper limits of the interquartile ranges of the EASE tool (HSE, 2003), and can therefore best be considered 75th percentiles (P75) of exposure distributions for particular ESs.

Between-user reliability is another major issue of the use of the exposure models. If two users select different model input parameters for the same exposure situation, this may lead to different exposure estimates and, more importantly, different judgement about the risk. This between-user variability is added to the uncertainties inherent in the model itself and could lead to an even larger variance in the new exposure distribution. As for the models' validity, several studies (Landberg *et al.*, 2015; Lamb *et al.*, 2017; Savic *et al.*, 2019) tried to quantify the influence of between-user variability on the exposure estimates in different models. The largest disagreement between individual users is found for activity-related parameters (Lamb *et al.*, 2017), such as the PROC in the TRAv3. A sensitivity study (Riedmann *et al.*, 2015) found these input

parameters to contribute significantly to the final exposure estimate. This could explain why differences in the estimates calculated by different users ranged within two or even more orders of magnitude in previous studies (Lamb *et al.*, 2017, Savic *et al.*, 2019). This emphasizes the importance of collaboration and knowledge and training in the exposure modelling process which would improve the quality of validation studies. The study of Lee *et al.* (2019a, b), for example, involved six experts from different institutes who were asked to code first independently the contextual data and then reach a mutual agreement about which input parameters best represent the considered ESs. Notably, such an approach does not guarantee to completely remove the between-user variance. Unlike some other studies (Kupczewska-Dobecka *et al.*, 2011; Ishii *et al.*, 2017), which either did not report their number or only a single assessor interpreted the data, the approach by Lee *et al.* (2019a, b) and in the ETEAM project should be expected to result in more reliable conclusions.

This study investigated the performance of ECETOC TRAv3 by pooling and, if considered necessary, re-coding the data that were already used in other research studies that had considered the TRA (Urbanus *et al.*, 2020). Utilizing this curated data set as the reference, the overall performance of the TRAv3 was investigated in a series of stratified analyses and via multiple linear regression to analyze which input parameters contributed most to the model's identified over- and underestimations.

Methodology

Exposure data

The occupational inhalation exposure data were obtained from previous studies that investigated the performance of the TRAv3 (TRA further in the text). This process was conducted by teams of two reviewers per data source. As described previously (Urbanus *et al.*, 2020), before further processing, those ESs containing less than six individual measurements were discarded so that in total, 129 ESs were retained, that is, 10 for solids and 119 for liquids. For example, the ETEAM project alone provided measurement data for 52 ESs, based on a copy of the project's database provided by the German Federal Institute for Occupational Hygiene and Medicine to ECETOC. Each ES represents a single line in the assembled database (see [Supplementary Material](#), available at *Annals of Work Exposures and Health* online). The technical quality of the data for each ES were systematically scored using the criteria in [Table 1](#), based on a similar system used in another study (Franken *et al.*, 2020). All data have to clearly represent personal exposure measurements covering a single activity or several similar activities during a whole shift

and have indications of the duration of exposure as well as the duration of the measurements. Exposures had to be occurring as part of normal routine operations, covering activities that could be assigned unequivocally to a single process category (PROC) as included in TRA.

For each exposure situation, the 75th percentile (P75) was calculated either directly from the geometric mean and standard deviation (GSD) if these were available or if all individual measurement results were available, or from the arithmetic mean and standard deviation of the data set.

Coding exposure situation descriptors for TRA input

As a next step, the necessary data were extracted by the review teams from the publications, their corresponding supplementary material, or underlying documentation to determine the inputs for the TRA parameters. For example, part of the material used in the ETEAM project originates from US NIOSH Health Hazard Evaluation reports which are readily available online from health hazard evaluations | NIOSH | CDC. In the cases where the required information could not be collected, the authors were directly contacted to provide the raw exposure information and the input parameters they had used to calculate the estimates.

Based on the information retrieved and/or received, a verification of the input parameters coded by the publications' authors was conducted independently of the original authors. Two reviewers performed the review independently to minimize individual bias. In particular, the chosen PROC, type of setting (industrial versus professional), and type of general ventilation (GV) were verified. For a number of the ESs, the reviewers had reason to disagree with the input parameter selection by the original authors. Each identified disagreement was discussed internally in the project team until consensus was achieved and, where justified, the original coding was corrected. The reason for disagreement was documented in the database.

Next, the TRA estimates were verified or corrected and then compared with the corresponding measurement data sets. Since the TRA nominally estimates the 75th percentile of the distribution, equation (1) was used to calculate this percentile for the measured exposure. In this equation, the *z*-score of 0.674 was used to calculate the given percentile.

$$P = GM \times GSD^z \quad \text{Eq. 1}$$

Statistics and visualization

All calculations and the visualization of the obtained results were conducted in Python [including its packages, that is, NumPy (Harris *et al.*, 2020), pandas (pandas development team, 2020), and Matplotlib (Hunter and Matplotlib, 2007)].

Table 1. Criteria for evaluating and assigning reliability scores to exposure data and input data for generating TRA estimates (based on Franken *et al.*, 2020).

Score	Adequacy assignment	General criteria	Examples
1	Adequate without restriction	Data of good technical and contextual adequacy is available	Completely documented measurement studies, performed with validated measurement methods (published by renowned institutes) and with all information on each data point in annexes. Full and unambiguous data to select TRA input parameter settings
2	Adequate with restrictions	Data of at least acceptable technical adequacy and information on contextual adequacy is available or can be evaluated based on expert judgement and reasonable assumptions	Well-documented measurement studies, performed with validated measurement methods (published by renowned institutes) or methods that resemble such methods closely and for which sufficient information on validity, accuracy, precision, and boundaries is available; sufficient description of context to either directly know the values for relevant factors or to make informed and justified expert judgement on a number of factors; activities may need to be categorized, based on descriptions, assumptions on scale and setting may need to be based on expert judgement, data on substance and product characteristics may need to be found in other sources or estimated
3	Useful as supporting evidence	Data of limited technical adequacy	Measurements with undocumented sampling techniques; statistical summaries of data (vapour pressure of measured substances, concentrations of substances in products, or largely different settings) that are not stratified; studies in which only the jobs of sampled workers are indicated without any indication of activities being sampled.
4	Not adequate	Data for which the technical adequacy cannot be evaluated or that are described too insufficiently to allow evaluation of several factors related to contextual adequacy	Studies in which the sampling method is not described (e.g. no reporting of whether respirable dust, inhalable dust, or total dust has been measured); the method for measuring solid/liquid aerosols is not described; studies in which no information is given on e.g. the use or no use of localized control measures, the concentration of measured substances in articles, the duration of activities within shift-based measurements, the containment of sources, etc

Regression

A regression model was established to estimate intercept (a), slope (b), and R^2 -score between the measured and the modelled (i.e. TRA) estimates. Since occupational exposure usually follows a lognormal distribution, the log transformation was applied on the 75th percentile calculated from the measurements (P75) and the modelled exposure estimate. These data points were plotted to illustrate how they follow the established regression line. R^2 -score, or the coefficient of determination, was determined to show how much variance in the P75s of the measurements the TRA could explain.

DeltaTRA

This parameter was calculated to aid the visualization of local trends between the measured and modelled exposure. As shown in equation (2), the $\text{delta}_{\text{TRA}}$ is calculated as the difference between logs of the modelled (TRA) and its corresponding measured value (y). While positive $\text{delta}_{\text{TRA}}$'s indicate overestimation, negative values indicate an underestimation of the measurements

by the model. Plotting the $\text{delta}_{\text{TRA}}$'s over a range of the measured exposure showed within which ranges the model tends to over- or underestimate.

$$\text{delta}_{\text{TRA}} = \log \text{TRA} - \log y \quad \text{Eq.2}$$

Mean absolute error

The mean of the absolute differences between the modelled and measured exposure in equation (3) defines another performance measure called mean absolute error (MAE) (Walther and Moore, 2005) (equation 3). While $\text{delta}_{\text{TRA}}$'s are calculated for all data points, MAE is calculated as a single value. This parameter shows how far, on average, the modelled estimates are away from the measured values for a data set with n ESS.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\log \text{TRA} - \log y| \quad \text{Eq.3}$$

If, for example, MAE equals 1.0, this would mean that the modelled and measured values differ on average by

one order of magnitude or a factor of ten since the difference is on the log scale.

Precision

This model's behavior shows how much the model is consistent in over- or under-estimating the exposure. For precision, we calculated how many modelled estimates were lying between different factors ϕ from the measurements. For example, we calculated the percentage of ESs for which the modelled estimates were higher than the measured P75 value by a factor between 5 and 10.

Delta_{TRA}'s per input parameter

Until this point of the analysis, the described steps used the dataset as a whole to investigate the behaviour of the TRA. However, it was also necessary to identify parameters (e.g. *ventilation*) and/or their categories (e.g. *GV* or *enhanced GV*) for which the TRA could underperform. Since only 10 data sets of six or more measurements and with sufficient data quality were available for solids, they were not analysed for each parameter. For ESs involving liquids, the dataset was split into smaller subsets for which delta_{TRA}'s (equation 2) were plotted. The split was performed separately for vapour pressure, volatility bands, PROCs, type of setting, ventilation, and the presence of LEV. For volatility bands, for example, the dataset (without solids) was split into three subsets that included only the ESs with low (0.01–500 Pa), medium (500–10000 Pa), or high (>10000 Pa) volatility substances. In this case, the plotted delta_{TRA}'s shows for which volatility band TRA estimates are most likely over- or underestimating the measurements.

Multiple linear regression

Equation (4) is the multiple linear regression model that was established to quantify the combined influence of the various input parameters on the difference between the modelled and measured values (i.e. on the delta_{TRA}'s; see equation 2). Again, this was not applied to the ESs for solids due to the small database.

$$\text{delta} = \omega_0 + \omega_{\text{vol}}\text{vol} + \omega_{\text{c}}\text{c} + \omega_{\text{proc}}\text{proc} + \omega_{\text{ts}}\text{ts} + \omega_{\text{vent}}\text{vent} + \omega_{\text{lev}}\text{lev} + \omega_{\text{t}}\text{t} \quad \text{Eq.4}$$

In equation (4), the influence on the delta_{TRA}'s (delta) was evaluated for volatility band (vol), concentration band in the product (c), PROCs (proc), type of setting (ts), room ventilation (vent), presence of LEV (lev), and task duration band (t). Since all these parameters are categorical values (e.g. PROC-7-*industrial spraying*), they were converted to dummy values. For each input parameter, the established linear regression evaluated a regression coefficient, ω . We used these coefficients to quantify how different parameter entries affect the

model's over- and underestimations. Within each input parameter, the multiple linear regression selected one category as a reference category and to which zero was assigned as its regression coefficient, ω . The multiple linear regression model selected the reference category using the alphabetical order within each input parameter. This is why, for example, *industrial* is a reference category and not a *professional setting* and similarly, PROC 10 acted as a reference category. The regression coefficients for all other categories (e.g. other PROCs) of a given input parameter were then evaluated relative to the reference category. This means that a category for which a positive coefficient is found is likely to result in a higher positive or smaller negative difference between the modelled and measured exposures than the reference category.

Results

Creation of the curated database

Table 2 summarizes the constructed database (available as [Supplementary Material](#), available at *Annals of Work Exposures and Health* online) and its main sources. The largest part of the database covers ESs for vapours from volatile liquids (119 ESs, 2171 measurements). A much smaller part covers ESs for dusts from solids (10 ESs, 101 measurements). Compared with the previous publication of this project ([Urbanus et al., 2020](#)) one further TRA performance study was added to the review which covered inhalation exposure measurements and TRA exposure modelling for 20 handling scenarios of petroleum substances ([Hesse et al., 2018](#); [ECETOC, 2022](#)). Data sets for 10 ESs (60 measurements) from the original materials were rated as not valid for the comparison exercise, in particular, because the measurements covered several worker activities that should be described by more than a single PROC, or the publication provided insufficient contextual information and, when followed up, the original authors had no longer access to their detailed measurement

Table 2. Main information sources.

Main sources	Number of exposure situations	Total number of measurements
ETEAM project	52	1045 (ART, NIOSH, SUVA)
CONCAWE	20	282
US NIOSH 2019	14	157
S. Lee (Korea)	18	103
Ishii (Japan)	7	72

notes. Details of the quality scoring are documented in ECETOC Technical Report 140 (ECETOC, 2022).

The review confirmed the majority of the original TRA input parameter selections but identified a number of required corrections. The percentage of corrections per parameter is indicated in Table 3. PROC selection was the parameter with the largest percentage (25) of corrections. Examples of these corrections included a change from PROC 2 to PROC 4, because the described activity was not a continuous process, a change from PROC 10 to PROC 7 for an activity described as spraying, and a change from PROC 13 to PROC 15 for activities with small quantities of substance (less than 1l), typical for laboratories. Room ventilation status was corrected in 16% of the ESs, in particular when available data on room size and air volume extracted by fans allowed the calculation of the actual number of air changes per hour (ACH) to align with the TRA definitions of basic (1–3 ACH), general (3–5 ACH) and enhanced ventilation (>5 ACH).

Regression and R^2 -score

Figure 1 illustrates the relation between the measurement P75 values and the modelled exposure to dust from solids and to vapour from liquids on logarithmic scales. The unity lines where the measurement P75 values equal the modelled values are added for illustration. Table 4 shows their corresponding regression coefficients and R^2 scores. The evaluated slope was close to 1.0 for liquids and 0.71 for solids, while the intercepts were found to be negative for both physical states, implying that on average the measured data were lower than the modelled estimates. Based on the evaluated R^2 score, the regression model explained somewhat better the variance in the measurement P75 values taken for liquids than for solids (36 versus 24%).

Delta_{TRA}'s, precision and MAE

The calculated delta_{TRA}'s for liquids (equation 2) in Fig. 2 indicates a tendency of the TRA to overestimate

lower measurement P75s and underestimate higher ones. Table 5 shows the overestimations that the majority of these, that is, 60% for solids and 47% for liquids, were lying within factors 1–5 from the measurement P75s. Smaller percentages of the data were highly overestimated, that is, 3% for liquids by a factor of >100. TRA underestimated 20% of the measurement P75s for both solids and liquids, respectively, also within a factor of 5. Finally, MAE, which was <1.0, shows that the modelled estimates and the measurement P75s were, on average, within the same order of magnitude.

Delta_{TRA}'s per variable

For the liquids, delta_{TRA}'s for several specific variables were plotted to examine trends, as described here below. Due to the small data size, this analysis was not conducted for solids. No trend can be seen with an increase in the vapour pressure in Fig. 3 within volatility bands. The delta_{TRA}'s, however, decrease for all volatility bands with the increase in the measured exposure. Among PROCs (Fig. 4), PROC-7-*industrial spraying*, PROC-10-*roller application or brushing*, and PROC-14-*production of preparations or articles by tableting, compression, extrusion, or pelletization* were more prone to underestimations than the other PROCs. Higher tendencies for underestimations were also found for *industrial* versus *professional* type-of-settings and for ESs with a LEV (LEV) measure present over those without. For GV, only for *enhanced (mechanical) ventilation* (air exchange rate of at least 5 per hour) there were no cases of underestimation, whilst the other options of no ventilation, GV and outdoor setting had various portions of underestimated ESs.

Multiple linear regression (MLR)

By splitting the data set repeatedly into training and test subsets and using different ratios between the two sets, the maximal R^2 score of 0.503 and an adjusted R^2 of 0.356 were obtained when the split was performed with an 80:20% ratio in favour of the former. This means that when the multiple regression model is trained, it can explain half of the variance in the test data. Shuffling the data or changing the ratio between the two subsets resulted in lower R^2 -score values, which were sometimes almost zero.

Table 6 summarizes the evaluated intercept and regression coefficients compared with the reference categories for the TRA's variables considered in the multiple regression model of the delta_{TRA}'s. Solids were again not processed due to insufficient data size. The lowest coefficients were obtained for ESs that included medium volatility, PROCs 7, 14, and 19, industrial settings, GV, with LEV present, and shorter durations,

Table 3. Percentage coding corrections applied to original materials for input parameters to generate TRA exposure predictions.

Parameter	Corrected (%)
PROC	25
Room ventilation status	16
Local exhaust ventilation	8
Setting (IND, PROF)	6
Application of duration factor	5
Concentration substance in product	3
Substance fugacity	1

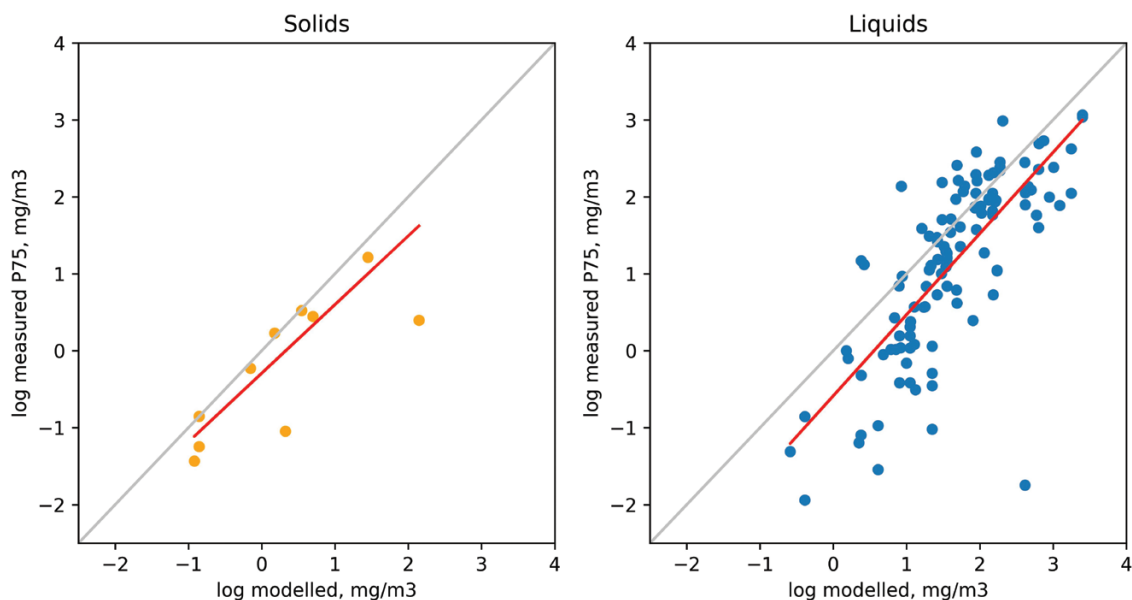


Figure 1. Measured 75th percentile (P75) versus modelled (TRA) exposure for solids and liquids. Both figures: thick line represents the regression line, thin line represents unity (P75 = TRA) added for illustration. For solids: $y = 0.7053 \times -0.3792$ ($R^2 = 0.64$). For liquids: $y = 1.0671 \times -0.6437$ ($R^2 = 0.6079$).

Table 4. Regression coefficients and R^2 -score.

	Intercept, a	Slope, b	R^2 -score
Solids	-0.38	0.71	0.24
Liquids	-0.65	1.07	0.36

meaning that for these input selections, the TRA is more likely to underestimate exposure.

Discussion

This study investigated the overall and per-input parameter performance of the ECETOC TRAv3 (TRA) by pooling and reusing existing exposure data from earlier TRA performance studies. Since the TRA is intended to provide a 75th percentile estimate of the exposure distribution for an activity, 75th percentiles (P75) calculated from data sets of six or more measurements for a given exposure situation were selected for analysis. They provide a more certain basis for comparison than individual measurement data as were done in several other projects, because some of these individual measurements could have been less representative of the workplace situation. The contextual data used in those previous studies were compared with their assigned input parameters in the TRA. Where a disagreement with these parameters was identified, a consensus-based correction was done by the authors of this study

and documented in a curated database. Every such modification required the exposure to be recalculated. Consequently, a number of the TRA exposure predictions obtained for this study differed from those in the original studies.

The results for solids were in general agreement with the ETEAM project. Overall, however, the TRA was found to be more conservative than suggested in both ETEAM and the study conducted by Lee *et al.* (2019b). The insufficient conservatism for solids could be due to the limited data that were available (i.e. 10 ESs versus 119 ESs for liquids) and would need to be confirmed when more data become available. An interesting result is that the modelled and measured exposures were mostly within the same order of magnitude. More precisely, we found that the TRA's estimates were distanced by not more than a factor of five from the measurements for 80% of the ESs for solids and 63% for liquids, including both over- and underestimations. While the majority of the underestimations found were thus within this range, for 20 and 25% ESs for solids and liquids, respectively, this difference was by 1–2 orders of magnitude.

The TRA's conservatism was investigated also graphically using Δ_{TRA} 's, that is, visualizing the difference between the modelled and measured exposures on a logarithmic scale. This approach is useful to investigate the model's conservatism within certain ranges and allows us to find a potential trend. Indeed, it was found that the TRA's conservatism decreases from lower to

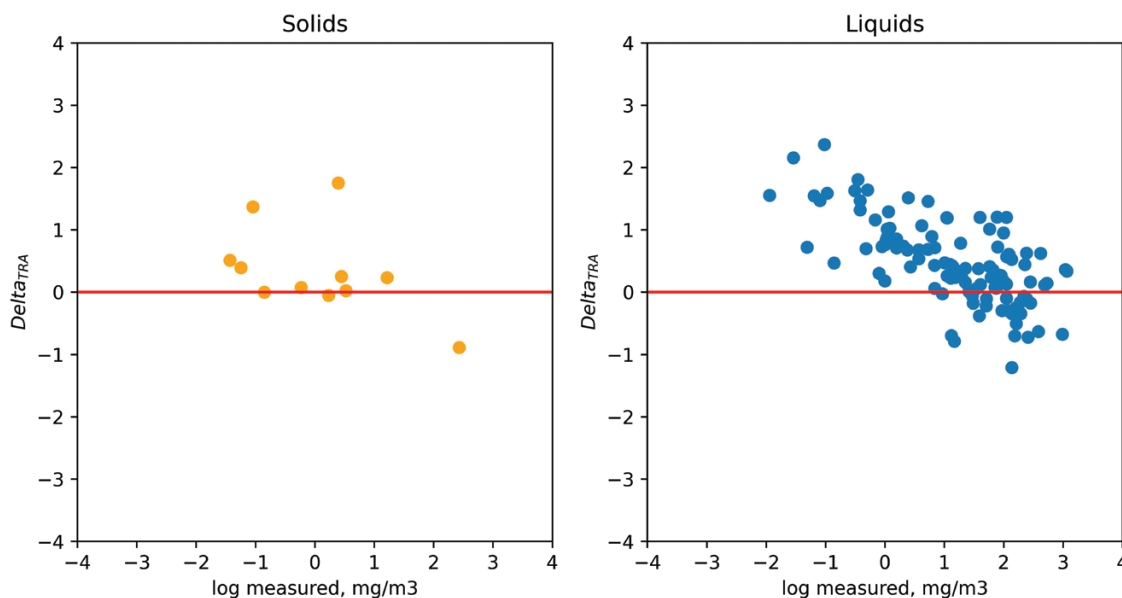


Figure 2. DeltaTRA's [log (P75) – log (TRA)] versus 75th percentiles of measurement data sets for solids and liquids.

Table 5. Precision and MAE. Precision is estimated as the percentage of the modelled values lying between factors ϕ from the measured value.

	ϕ								MAE
	-1000 to (-100)	-100 to (-10)	-10 to (-5)	-5 to (-1)	1-5	5-10	10-100	100-1000	
Solids	0	0	0	20	60	0	20	0	0.5
Liquids	0	1	3	16	47	8	22	3	0.7

higher measurements. Previous publications (Savic *et al.*, 2017; van Tongeren *et al.*, 2017, Lee *et al.*, 2019b) reported similar findings for several models including the TRA (Lee *et al.*, 2019b). A consequence is that the model tends to overestimate lower and underestimate higher measurements.

The analysis results obtained for individual input parameters and the established multiple regression coefficients explained which input parameter underestimations are more likely. Figures 3 and 4 illustrate the results graphically, subdivided for various TRA input parameters. This study found that TRA predictions with *low volatility*, *professional setting*, and ESS without LEV are more conservative compared to their alternatives (e.g. *industrial setting*). As in the ETEAM study, PROC-7, PROC-14, and PROC-19 were more prone to underestimation than the other process categories. Underestimations in our case were also obtained for PROC-10. On the other hand, our results for PROC-15 were more numerous and showed a tendency to overestimate, contrary to the ETEAM study. These results were further analysed using a multiple

regression model because all available categories defined for the input parameters were not uniformly distributed in all data sets. For example, PROC-7 only addresses *industrial settings*. To answer the question of whether the underestimations encountered were caused by PROC or other types of input parameters a multiple regression model was therefore established to investigate their independent effects on the degree of under- or overestimations. By combining the evaluated regression coefficients and the intercept in Table 6, the expected behaviour for a given set of input parameters can then be calculated. Equation (5) shows an example of such a calculation for the following set of input parameters:

- Volatility: *low*
- Concentration > 25%
- PROC: 10
- Type of settings: professional
- Ventilation: GV
- LEV: *no*
- Duration > 4h

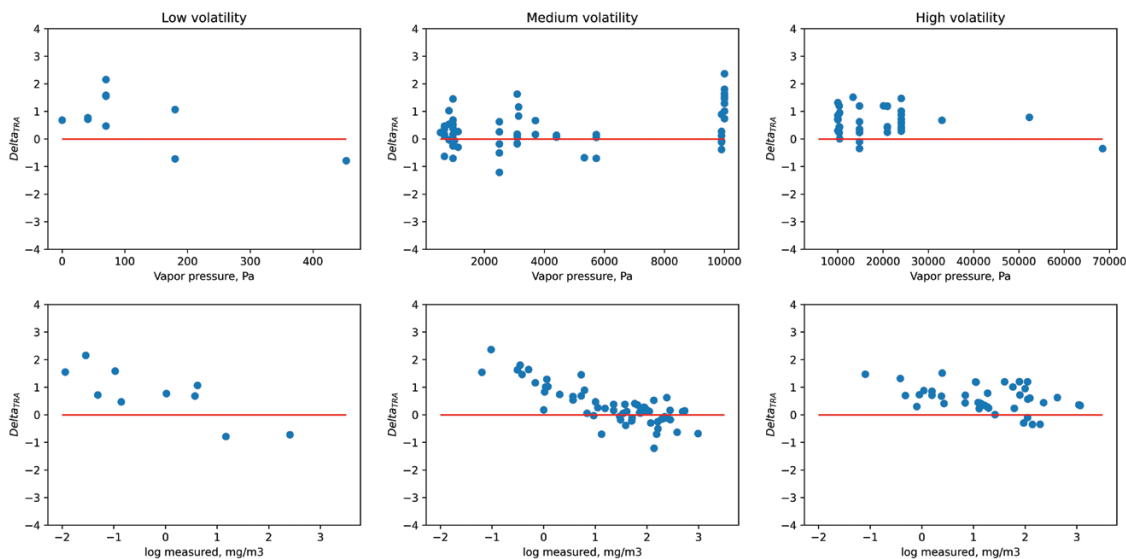


Figure 3. DeltaTRA's [log (P75) – log (TRA)] versus vapour pressure (top row) and 75th percentiles of measurement data sets (bottom row) for the three volatility bands (liquids only).

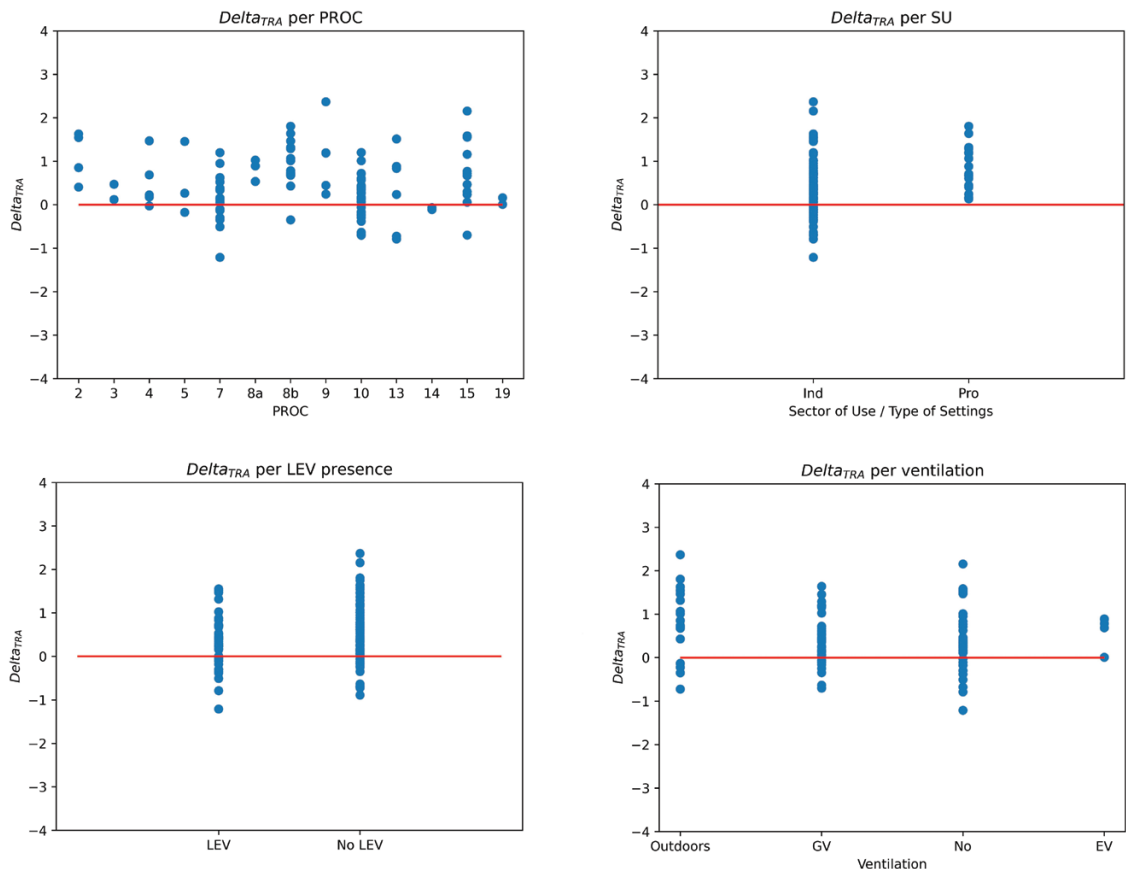


Figure 4. DeltaTRA's [log (P75) – log (TRA)] per process category (PROC), sector of use (SU)/type-of-setting (IND: industrial, PRO: professional), presence of LEV and room ventilation [GV: general ventilation (3–5 air changes per hour), EV: enhanced ventilation (5–10 air changes per hour)] (liquids only).

Table 6. Summary of multiple regression coefficients obtained for liquids (Δ_{TRA} 's).

Input parameter	Category	Coefficient	$P > t $	CI (25th–75th)
–	Intercept, a	0.65	0.13	–0.20 1.49
Volatility band	Low	–0.29	0.23	–0.77 0.18
	Medium	–0.72	0.12	–1.64 0.19
	High	0.00		
Concentration	> 25%	–0.03	0.88	–0.42 0.36
	5–25%	0.29	0.23	–0.18 0.76
	1–5%	0.00		
	< 1%	1.06	0.02	0.15 1.97
PROC	PROC-2	0.58	0.09	–0.09 1.25
	PROC-3	0.05	0.90	–0.77 0.88
	PROC-4	0.02	0.95	–0.61 0.65
	PROC-5	0.26	0.35	–0.30 0.83
	PROC-7	–0.09	0.62	–0.43 0.26
	PROC-8a	0.54	0.11	–0.12 1.20
	PROC-8b	0.18	0.40	–0.25 0.62
	PROC-9	0.45	0.21	–0.26 1.15
	PROC-10	0.00		
	PROC-13	0.62	0.07	–0.04 1.28
Type of settings	PROC-14	–0.28	0.50	–1.10 0.54
	PROC-15	0.65	0.01	0.17 1.15
	PROC-16	0.44	0.48	–0.80 1.68
	PROC-19	–0.57	0.22	–1.48 0.34
	Industrial	0.00		
	Professional	0.35	0.06	–0.02 0.73
	Ventilation	No general ventilation	–0.14	0.66
General ventilation (GV)		–0.25	0.44	–0.88 0.38
Enhanced GV		0.00		
Local Exhaust Ventilation (LEV)	Outdoors	0.03	0.94	–0.71 0.76
	No LEV	0.00		
	LEV present	–0.33	0.03	–0.63 –0.03
Duration	> 4 h	0.15	0.34	–0.17 0.47
	1–4 h	0.00		
	15 min–1 h	–0.35	0.26	–0.97 0.27
	< 15 min	–0.51	0.18	–1.26 0.23

$$\Delta = 0.65 - 0.29 - 0.03 + 0.00 + 0.35 - 0.25 + 0.00 + 0.15 \quad \text{Eq.5}$$

In equation (5), the first value is the intercept from Table 4. The obtained result equals 0.58, which means that for the provided set of input parameters and based on the data in our curated database, an overestimation by the TRA of the measured exposure by a factor of $10^{0.58}$ or 3.8 is predicted; the modelled and measured values are thus within the same order of magnitude. In other words, all categories in Table 6, for which

a positive coefficient was evaluated, are in favour of overestimations. If the outcome of equation 5 for a particular ES is a negative value, then underestimation is expected. Regarding PROCs, whilst the majority have a positive coefficient compared to PROC 10, the lowest coefficients were found for PROC-7 (–0.09), PROC-14 (–0.28), and PROC-19 (–0.57). A higher probability for underestimation can be expected also for *medium volatility*, *industrial settings*, *GV*, and with *LEV present*. It is obvious that these results are similar to those illustrated in Figs. 3 and 4, but the regression results are more precise as they quantify the likelihood

of underestimations for a given category versus its counter pair(s) for the input parameter. However, very few coefficients were statistically significant.

One input parameter available in the TRA could not be examined in the current study, i.e. respiratory protective equipment (RPE). Occupational exposure measurement is conventionally carried out in the worker's breathing zone outside any RPE that is worn. In the case of RPE use the measurement result does therefore not directly reflect the composition of the inhaled air. The protection afforded by RPE has been adopted in the TRA based on externally available references which have not been questioned in the TRA validation studies. When calculating over- and underestimation of the TRA using the MLR equation for a given exposure scenario with RPE use, the RPE parameter should therefore be first excluded from the modelled estimate.

It should be noted that despite the large effort on data review, correction, and pooling, the available database is not complete for all possible TRA scenarios, and thus all potential combinations of the input parameters' categories. In addition, the number of data points per scenario is also not balanced, hence the conclusions presented on overall TRA performance and any recommendations for improvement of TRA remain conditional. Nevertheless, the available data and conducted analysis allowed us to give more weight to the detailed results per scenario and on individual parameters than to overall tool performance indicators presented by previous studies. For example, no instances of underestimation of Professional scenarios were found and many other scenarios were found to be predicted in a suitably conservative way. Based on the stratified analysis results, several tool settings were identified that will be corrected in a future version of the TRA, including LEV efficiency which will be reduced from 95 to 90% for PROC-7 and PROC 8b/industrial, and from 90% to 80% for PROC 8b/professional. These changes will address in part the observed lack of conservatism in the LEV parameter, as well as better align this parameter across the tool. Also, the base estimates for medium volatility liquids for PROC 10 will be doubled. The scenarios involving these parameter settings contributed a significant portion of the datasets to the current overall curated database with a consequent impact on the observed overall tool performance. The planned corrections will reduce the underestimations in the curated database from 20 to 13% (data not shown). Other apparent underestimations (e.g. scenarios with PROC 14 and 19, see also Fig. 4) had insufficient observations to identify a suitable correction. With these targeted adjustments, the TRA can continue to fulfil its role as an efficient tier-1 or screening tool in REACH chemical safety assessment of occupational exposure to chemicals. Further

enhancement of the validation database is desirable and should preferably focus on data gaps, not on those scenarios for which conclusions have already been reached.

Acknowledgements

The authors are grateful to Professor David Vernez (Center for primary care and public health (Unisanté), University of Lausanne, Route de la Corniche 2, 1066 Epalinges-Lausanne, Switzerland) for careful review of a draft of the manuscript and recommendations for completion. The authors are also very grateful to Francesca Borghi and Andrea Spinazzè (Università degli Studi dell'Insubria, Como, Italy), Urs Schlüter (German Federal Institute for Occupational Safety and Health) and Wouter ter Burg (Dutch National Institute for Public Health and Environment) who acted as external advisory panel members throughout this project. Finally, the authors are grateful to Andreea Cuciureanu, Consultant PFA-Brussels SPRL for ECETOC, for overall project coordination.

Conflict of interest

All listed authors are members of the ECETOC TRA-Worker task force. ECETOC funded Nenad Savic' contribution to the project. All other authors contributed in the course of their regular employment or without funding.

Data availability

The data underlying this article are available in summary form in the article and in its online supplementary material, as well as in ECETOC Technical Report 140 (ECETOC 2022).

Supplementary data

Supplementary data are available at *Annals of Work Exposures and Health* online.

References

- ECETOC. (2004) *Technical Report 93 - Targeted Risk Assessment*. ECETOC: Brussels, Belgium.
- ECETOC. (2009) *Technical Report 107 - Addendum to ECETOC Targeted Risk Assessment Report No. 93*. ECETOC: Brussels, Belgium.
- ECETOC. (2012) *Technical Report 114 - ECETOC TRA version 3: Background and Rationale for the Improvements*. ECETOC: Brussels, Belgium.
- ECETOC. (2022) *Technical Report 140 - Systematic Review of Published Studies of ECETOC TRA Worker Exposure Predictions*. ECETOC: Brussels, Belgium.

- ECHA. (2016) *Guidance on information requirements and chemical safety assessment. Chapter R.14: Occupational exposure assessment*. Version 3.0, August 2016. European Chemicals Agency, Helsinki, Finland.
- Franken R, Shandilya N, Marquart H *et al.* (2020) Extrapolating the applicability of measurement data on worker inhalation exposure to chemical substances. *Ann Work Expo Health*; **64**: 250–69.
- Fransman W. (2013) *Development of a mechanistic model for the Advanced REACH Tool (ART)—Version 1.5—(updates previous version 1.0, June 2010)*. Zeist, The Netherlands: TNO.
- Harris CR, Millman KJ, van der Walt SJ *et al.* (2020) Array programming with NumPy. *Nature*; **585**: 357–62.
- Hesse S, Hahn S, Lamb J *et al.* (2018) *Review of Tier 1 Workplace Exposure Estimates for Petroleum Substances in REACH dossiers*, Brussels, Belgium: CONCAWE.
- HSE. (2003) *Evaluation and further development of the EASE Model 2.0*. Norwich, UK: Health and Safety Executive Books.
- Hunter JD. (2007) Matplotlib: a 2D graphics environment. *Comput Sci Eng*; **9**: 90–5.
- Ishii S, Katagiri R, Kitamura K *et al.* (2017) Evaluation of the ECETOC TRA model for workplace inhalation exposure to ethylbenzene in Japan. *J Chem Health Safety*; **24**: 8–20.
- Kindler P, Winteler R. (2010) *Anwendbarkeit von Expositionsmoellen für Chemikalien auf Schweizer Verhältnisse. Teilprojekt 1: Überprüfung der Modelle "EASE" und "EMKG-Expo-Tool"*.
- Kupczewska-Dobecka M, Czerczak S, Jakubowski M. (2011) Evaluation of the TRA ECETOC model for inhalation workplace exposure to different organic solvents for selected process categories. *Int J Occup Med Environ Health*; **24**: 208–17.
- Lamb J, Galea KS, Miller BG *et al.* (2017) Between-user reliability of tier 1 exposure assessment tools used under REACH. *Ann Work Expo Health*; **61**: 939–53.
- Landberg HE, Berg P, Andersson L *et al.* (2015) Comparison and evaluation of multiple users' usage of the exposure and risk tool: stoffenmanager 5.1. *Ann Occup Hyg*; **59**: 821–35.
- Lee EG, Lamb J, Savic N *et al.* (2019a) Evaluation of exposure assessment tools under REACH: part II-higher tier tools. *Ann Work Expo Health*; **63**: 230–41.
- Lee EG, Lamb J, Savic N *et al.* (2019b) Evaluation of exposure assessment tools under REACH: part I-tier 1 tools. *Ann Work Expo Health*; **63**: 218–29.
- Marquart H, Heussen H, Lefeber M *et al.* (2008) 'Stoffenmanager', a web-based control banding tool using an exposure process model. *Ann Occup Hyg*; **52**: 429–41.
- Mc Donnell PE, Schinkel JM, Coggins MA *et al.* (2011) Validation of the inhalable dust algorithm of the Advanced REACH Tool using a dataset from the pharmaceutical industry. *J Environ Monit*; **13**: 1597–606.
- Money C, Schnoeder F, Noij D *et al.* (2014) ECETOC TRA version 3: capturing and consolidating the experiences of REACH. *Environ Sci Process Impacts*; **16**: 970–7.
- pandas development team, (2020) T., Pandas. Zenodo. doi:10.5281/zenodo.3509134
- Riedmann RA, Gasic B, Vernez D. (2015) Sensitivity analysis, dominant factors, and robustness of the ECETOC TRA v3, Stoffenmanager 4.5, and ART 1.5 occupational exposure models. *Risk Anal*; **35**: 211–25.
- Savic N, Gasic B, Schinkel J *et al.* (2017) Comparing the advanced REACH Tool's (ART) estimates with Switzerland's occupational exposure data. *Ann Work Expo Health*; **61**: 954–64.
- Savic N, Lee EG, Gasic B *et al.* (2019) Inter-assessor agreement for TREXMO and its models outside the translation framework. *Ann Work Expo Health*; **63**: 814–20.
- Schinkel J, Heussen H, Fransman W *et al.* (2010) Cross-validation and refinement of the Stoffenmanager as a first tier exposure assessment tool for REACH. *Occup Environ Med*; **67**: 125–32.
- Schinkel J, Warren N, Fransman W *et al.* (2011) Advanced REACH Tool (ART): calibration of the mechanistic model. *J Environ Monit*; **13**: 1374–82.
- Urbanus J, Henschel O, Li Q *et al.* (2020) The ECETOC-targeted risk assessment tool for worker exposure estimation in reach registration dossiers of chemical substances-current developmentS. *Int J Environ Res Public Health*; **17**: 8443.
- van Tongeren M, Lamb J, Cherrie JW *et al.* (2017) Validation of lower tier exposure tools used for REACH: comparison of tools estimates with available exposure measurements. *Ann Work Expo Health*; **61**: 921–38.
- Walther BA, Moore JL. (2005) The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*; **28**: 815–29.