Author Manuscript

Faculty of Biology and Medicine Publication

This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Published in final edited form as:

Title: Evaluation of Exposure Assessment Tools under REACH: Part II-Higher Tier Tools.

Authors: Lee EG, Lamb J, Savic N, Basinas I, Gasic B, Jung C, Kashon ML, Kim J, Tischer M, van Tongeren M, Vernez D, Harper M

Journal: Annals of Work Exposures and Health

Year: 2019 Feb 16

Issue: 63

Volume: 2

Pages: 230-241

DOI: 10.1093/annweh/wxy098

Evaluation of exposure assessment tools under REACH: Part II—Higher Tier tools

Eun Gyung (Emily) Lee,[1]* Judith Lamb,[2] Nenad Savic,[3] Ioannis Basinas,[2] Bojan Gasic,[4] Christian Jung,[5] Michael L. Kashon,[6] Jongwoon Kim,[7] Martin Tischer,[5] Martie van Tongeren,[2,8] David Vernez,[3] Martin Harper[1]

[1]National Institute for Occupational Safety and Health (NIOSH), Health Effects Laboratory Division (HELD), Exposure Assessment Branch (EAB), 1095 Willowdale Road, Morgantown, WV 26505, USA
[2]Institute of Occupational Medicine (IOM), Avenue North, Heriot Watt Research Park, Riccarton, Edinburgh, EH14 4AP, UK
[3]Institute for Work and Health (IST), University of Lausanne and Geneva, 1066 Epalinges-Lausanne, Switzerland
[4]Swiss State Secretariat for Economic Affairs (SECO), Bern, Switzerland
[5]Federal Institute for Occupational Safety and Health (BAuA), Dortmund, Germany
[6]NIOSH, HELD, Biostatistics and Epidemiology Branch, 1095 Willowdale Road, Morgantown, WV 26505, USA
[7]Korea Institute of Science and Technology (KIST) Europe, Campus E 7.1, D-66123 Saarbruecken, Germany
[8] Centre for Occupational and Environmental Health, Centre for Epidemiology, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK

*Author to whom correspondence should be addressed (Tel: 1-304-285-6041, email: dtq5@cdc.gov)

Keywords: Exposure Assessment Tools, REACH Tier tools, Inhalation tools, exposure assessment tools, inhalation tools, exposure assessment tool validation, REACH higher tools

ABSTRACT

The Stoffenmanager®v4.5 and Advanced REACH Tool (ART), two higher tier exposure assessment tools for use under REACH were evaluated,. A total of 282 exposure measurements from 51 exposure situations (ESs) were utilized and results were presented by exposure category, handling/activity description, and input parameters. For all exposure categories except for powders and solid objects, the Stoffenmanager appeared to be accurate (relative bias ranging from -56 to 332%) and robust to predict exposures (percent of measurements exceeding the tool's 90^{th} percentile estimate [%M>T] ranging from 0 to 15%). Yet, areas that this tool need to be improved were present including the handling of liquids on large surfaces or large work pieces task, the high and medium allocation of vapour pressure input, and the absence of local exhaust ventilation input. For handling solid objects, the Stoffenmanager assumed minimal emissions leading to zero exposure estimates, whereas low exposure levels were observed from field surveys; it is therefore recommended that the tool developer review an algorithm for this category. Although the ART's predictions appeared to be accurate for all exposure categories except powders, the %M>Ts for the 90^{th} percentile estimates were greater than 25%, indicating that the tool considerably underestimates exposures. Hence, it is strongly suggested that ART developers review the tool's calibration and assigned scores representing exposure variability towards improving the tool performance. Except for liquids with VP > 10 Pa, the findings of other exposure categories were based ona limited number of exposure situations and measurements with low ranges of concentrations. In addition, for both tools, not all handling/activity descriptions and input parameters were considered. Thus, further validation studies are still necessary.

INTRODUCTION

Tier 1 exposure assessment tools described in the European Chemicals Agency (ECHA) R14 guidance (2016) are designed to be simple and easy to use and to be conservative by overestimating a potential exposure for a defined exposure scenario. If exposures estimated from the first tier tools exceed the DNEL of a substance, it is recommended to use higher tier tools which are developed to generate exposure levels with greater accuracy and less uncertainty. Higher tier tools include Stoffenmanager® (hereinafter referred to as Stoffenmanager) and the Advanced REACH Tool (ART). The background information of both tools are well described in the ECHA R14 guidance and previous publications (Fransman et al., 2011; McNally et al., 2014; Marquart et al., 2008; Schinkel et al., 2011; Schinkel et al., 2013; Schinkel et al., 2014; Tielemans et al., 2008; Tielemans et al., 2011). In addition, a conceptual evaluation for the Stoffenmanager has been performed by Hesse et al. (2015).

Currently, few studies are available regarding the external validity of the tools (Annex A of Part I paper by Lee et al., 2017). Schinkel et al. (2010) validated the Stoffenmanager and suggested a refinement of the tool. As results, they included the datasets used for the validation study and developed four different equations for the following categories: (1) handling powders and granules; (2) handling which results in comminution, (3) handling low-volatile chemicals, and (4) handling volatile chemicals. Koppisch et al. (2012) validated equations of the Stoffenmanager regarding handling of powders/granules (n=390) and machining (n=1133) for exposures to inhalable dust. The authors used data from the MEGA ("Measurement data relating to workplace exposure to hazardous substances"; in German, "Messdaten zur Exposition gegenüber Gefahrstoffen am Arbeitsplatz") database, collected between 2000 and 2009, and reported 11% and 7% of the exposure data to be above the estimated 90th percentile of the Stoffenmanager for handling of powders and machining, respectively. Vink et al. (2010) evaluated the Stoffenmanager v4 with exposure data collected from paint spraying tasks and reported less conservative results compared to those for estimates by the ECETOC TRAv2. In a more recent study, Landberg et al. compared the 90th percentile estimates of Stoffenmanager v5.1 with the median values of exposure measurements for 11 distinct exposure scenarios across the wood, printing, metal foundry and spray painting industries. The authors reported the tool estimates to overestimate exposure for solids (n=11) but to underestimate exposure for liquids (n=26) by an approximate 27% . McDonnell et al. (2011) compared the exposure data (n=190) from 16 exposure scenarios collected in pharmaceutical companies with the ART predictions estimated from a refined version for the inhalable dust algorithm and reported the tool to underestimate the geometric mean exposure levels by approximately one-third. Recently, Spinazze et al. (2017) evaluated the accuracy and robustness of the Stoffenmanager v.6 and ART v1.5 using pre-existing exposure data of organic solvents and pesticides. As results, they reported that overall the ART was the most accurate for both organic solvents and pesticides but underestimated the exposures to pesticides. In terms of robustness, they concluded that the Stoffenmanager was the most robust tool showing less percent of measurements exceeding the model's 90th percentile estimates.

The present study has been conducted as a follow up on the The Evaluation of Tier 1 Exposure Assessment Models used under REACH (ETEAM) project to compare and evaluate the different REACH exposure tools using measurement data collected specifically for the purpose (Tischer et al., In Preparation). The results of the validations for the tier one tools are reported in Part I (Lee et al 2017) whereas this paper describes the results of the validation of the higher tier tools (Stoffenmanager v4.5 and ART) using the same exposure data.

METHODS

Field surveys and development of exposure situation (ES)

The field surveys and development of ES scenarios is described in detailed in Part I (Lee et al., 2017). Five exposure categories were defined as follows: (1) Aqueous solutions, (2) Liquids with a vapour pressure (VP) ≤ 10 Pa at room temperature, (3) Liquids with a VP > 10 Pa at room temperature, (4) Powders, and (5) Solid objects. An exposure category of metal processing was not included in this part because Stoffenmanager and ART have not been calibrated for hot metal processing tasks (ES9-Casting task and ES15-Smelting task). As a result, a total of 282 personal exposure measurements from 51 ESs were included (Table 1). Detailed information about tasks and number of samples for each task are listed in Supplement Table S1.

Although it is possible to consider multiple tasks in Stoffenmanager, this does not apply for the free web-version (v4.5). In this study, therefore, we combined several sub-tasks into one task and applied the lowest control method, same as Tier 1 tools. On the other hand, the ART is capable of accounting for a maximum of four sub-tasks.

Translation of contextual information into the tools' input parameters

The translation of contextual information into the tool input parameters was done concurrently to the translation for Tier 1 tools by the same assessors (EL, JL, NS, BG, JK, and MT) from six different organizations (Lee et al., 2017)For Stoffenmanager, assessors were asked to use the Microsoft Access database, developed as part of the ETEAM project. For ART, due to the complexity of the tool mechanism, each assessor was asked to create an account in ART (**Fehler! Hyperlink-Referenz ungültig.**) and to code the input parameters directly into the web-based tool. The ART allows the inputs and outputs to be exported into a Microsoft Excel report and the assessors were asked to send this report together with a summary of the range of ART-generated different percentiles of the predicted exposure and associated confidence intervals to the IOM and NIOSH. All the collected input parameters from the assessors were discussed at an in-present meeting in July, 2015 to make consensus of final inputs.

Generation of the tool estimates

With the agreed input parameters for each tool, the Stoffenmanager semi-quantitative scores were calculated using the published algorithms incorporated in the Access database. The score from the tool was then converted to a quantitative exposure estimation using the equations by Schinkel et al. (2010). For Stoffenmanager, the 50th, 75th, and 90th percentile estimates were reported. For ART, estimates were generated using the web-based tool and reported full-shift exposures of 50th, 75th, and 90th percentiles along with a 90% confidence interval of each percentile estimate.

Data analyses

Stoffenmanager is known as a task-based tool. Thus, as descripted in Part I study, it is necessary to modify the collected exposure measurements accordingly. We converted the exposure measurements with sampling times to task-based measurements and performed comparison with the tool's estimates. Stoffenmanager generates exposure estimates as a distribution form and a user can select specific percentiles. For example, if a user selects 90th percentile estimation, it means that the 90% of the exposure measurements in the exposure distribution with actual measurements were below than the selected 90th percentile estimation.

ART also provides exposure estimates in the form of a distribution but either for full-shift (recommended for REACH evaluation) or long-term exposures. A user can then select percentiles of the exposure distribution (i.e., 50th, 75th, 90th, 95th or 99th) and a confidence interval (CI) for each percentile (i.e., 80%, 90%, 95% or 99%) to be estimated. The former expresses exposure variability while the latter indicates the uncertainty around the percentile estimate. For the ART evaluation, all exposure measurements were converted into the full-shift exposures (i.e., 8-hour shift).

For both tools, the 90th percentile estimate was selected for the purpose of the evaluation, since it represents a "reasonable worst case" exposure situation and is recommended for risk characterization according to the ECHA R14 guidance (2016). In addition, 50th and 75th percentile estimates were obtained to compare the distribution of tools' estimates.

One way to determine whether the tool has been calibrated correctly is to compare accuracies of different percentiles (e.g., 50th, 75th, and 90th percentile) by calculating each percentile estimate from the exposure measurement distribution of each exposure situation. However, such a comparison could not be performed within the present study because of insufficient exposure measurements per exposure situation. Instead, we used 50th percentile estimates of both tools to calculate bias and precision using the following equations (Hornung, 1991):

$$Bias = \sum_{i=1}^{n_0} \frac{(\hat{y}_i - y_i)}{n_0} \quad , \qquad \Pr ecision = \sqrt{\sum_{i=1}^{n_0} \frac{\left[(\hat{y}_i - y_i) - bias\right]^2}{n_0 - 1}}$$

Relative bias was then calculated as (Schinkel et al., 2010):

$$Relative\ bias = \left(e^{bias} - 1\right) * 100\%$$

where $\hat{y}_i$ =predicted exposure level for the ith set of exposure factors in the validation set (log scale), y_i = measured exposure for the ith set of exposure factors (log scale), and n_0 = number of measurements in the validation set. The bias indicates a distance of the tool estimate from the true value, whereas the precision estimates variability. A positive bias implies overestimation by a tool compared to an exposure measurement; a negative bias implies underestimation. The smaller value of the relative bias indicates the more accurate results for the exposure estimation. Pearson correlation coefficient (r_p) was calculated to determine the relationship between exposure measurements and tool estimates (both log-transformed). All data analyses were performed using the Statistical Analysis Software (SAS) v. 9.4.

The results were presented by exposure category, handling/activity description, and tool input parameters. In situations where multiple activity descriptions were available for the ART, the activity description showing the longest exposure time was selected to represent the activity for a task. When two different activities had the same exposure time, these activities were combined, such as HC&A = Handling of contaminated objects or paste & Activities with relatively undisturbed surfaces (no aerosol formation).

Among 51 ESs, 38 ESs (~ 75%) included multiple sub-tasks with different control methods for the ART, but only the lowest control method was considered for the Stoffenmanager. For example, ES 19 (Batch-making task) had four sub-tasks with control methods of no presence of local exhaust ventilation and a fully enclosed system. For this ES, we considered the lowest control method (i.e., LEV absence) for the Stoffenmanager, whilst all control methods for each sub-task were applied for the ART. For these ESs, a direct comparison of accuracy between these tools would not be possible. Hence, the remaining 13 ESs (n=53) having only one task were pooled out from the full dataset to compare both tools in terms of accuracy.

RESULTS

Description of workplace measurement data

Tables 1 and 2 are summary of personal exposure measurements and tools' 90th percentile estimates presented by exposure category. Among 282 exposure measurements, the majority of measurements (~89%) falls in exposures to liquids with VP > 10 Pa. This exposure category includes a wide range of exposure concentrations from 0.01 mg/m^3 to 1455.4 mg/m^3 (full-shift measurements) and from 0.07 mg/m^3 to 6653.3 mg/m^3 (task-based measurements) covering workplaces from small laboratories to heavy industries. Compared to this category, other exposure categories showed lower ranges of exposure measurements. The number of ESs for liquids with VP > 10 Pa was 42, whereas the number of ESs for other exposure categories was ≤ 3.

Table 1. Summary of the personal exposure measurements (task-based) and the Stoffenmanager 90th percentile estimates (by exposure category)

Exposure Category	ES No	N	Exposure measurements				Stoffenmanager 90th percentile estimates			
			AM (mg/m^3)	GM (mg/m^3)	GSD	Range (mg/m^3)	AM (mg/m^3)	GM (mg/m^3)	GSD	Range (mg/m^3)
Aqueous solutions	2	4	0.92	0.73	2.31	0.24-1.82	31.87	28.95	1.67	18.5-45.2
Liquids with VP ≤ 10 Pa	2	5	0.07	0.05	2.83	0.02-0.16	0.35	0.31	1.83	0.16-0.48
Liquids with VP > 10 Pa	42	251	214.4	24.48	9.84	0.07-6653.3	459.48	284.92	2.72	30.5-1627.5
Powders	2	11	0.44	0.24	3.73	0.03-1.44	47.28	47.21	1.06	45.3-50.8
Solid Objects	3	11	0.09	0.03	5.64	0.002-0.33	0*	*	*	0*
Overall	51	282	190.9	13.40	16.78	0.002-6653.3	411.27	182.76	5.62	0.16-1627.5

Abbreviations: ES No=number of exposure situations (ESs) developed by NIOSH; N=number of personal exposure measurements; AM = arithmetic mean exposure; GM=geometric mean exposure; GSD=geometric standard deviation. * all estimates for this category are zero.

Table 2. Summary of the personal exposure measurements (full-shift) and the ART 90th percentile estimates (by exposure category)

Exposure Category	ES No	N	Exposure measurements				ART 90th percentile estimates			
			AM (mg/m^3)	GM (mg/m^3)	GSD	Range (mg/m^3)	AM (mg/m^3)	GM (mg/m^3)	GSD	Range (mg/m^3)
Aqueous solutions	2	4	0.15	0.07	5.26	0.01-0.37	0.06	0.0412	2.78	0.017-0.1
Liquids with VP ≤ 10 Pa	2	5	0.005	0.003	3.16	0.001-0.01	0.003	0.002	3.15	0.001-0.005
Liquids with VP > 10 Pa	42	251	60.9	7.88	10.30	0.01-1455.4	66.57	11.68	16.03	0.013-310
Powders	2	11	0.31	0.17	3.74	0.02-0.99	28*	*	*	28*
Solid Objects	3	11	0.03	0.02	4.24	0.001-0.083	0.335	0.040	8.95	0.011-1.2
Overall	51	282	54.2	4.32	17.49	0.001-1455.4	60.36	7.67	23.42	0.001-310

Abbreviations: ES No=number of exposure situations (ESs) developed by NIOSH; n=number of personal exposure measurements; AM = arithmetic mean exposure; GM=geometric mean exposure; GSD=geometric standard deviation. * The same estimate value was obtained for all 2 ESs.

Comparison of exposure measurements with tool estimates

1) **By exposure Category**

Both AMs and GMs of the 90th percentile Stoffenmanager predictions were higher than those of the measurements for all exposure categories except for the solid object category (Table 1). For this category the tiool predicted zero exposures for all 3 ESs included (ES 8 – Wire extraction, ES 10 – Packing and shipping, and ES 12 – Bar feeding). Correlations between the log-transformed Stoffenmanager estimates and log-transformed exposure measurements appeared to be high for exposures to liquids with VP ≤ 10 Pa, moderate for exposures to liquids with VP > 10 Pa, negative for aqueous solutions, and weak for the powder handling category. Note that Table 3 includes no results for the solid objects category because of zero exposure predictions for all ESs. Positive biases for the categories of aqueous solutions, liquids with VP > 10Pa, and powders indicate that on average the tool overestimated exposures when comparing with the measured data. In contrast,
aA negative bias for liquids with VP ≤ 10 Pa indicates underestimation of exposures. Among four categories, the tool showed the highest accuracy for exposures to liquids with VP > 10 Pa (relative bias =29%) and the lowest for exposures to powders (relative bias = 2003%). The percentage of measurements exceeding the tool's 90th percentile estimates were 15% for exposures to liquids with VP > 10 Pa category (Table 3 and Figure 1), whereas the other categories showed no measurements exceeding the tool estimates.

For the ART, all AMs and GMs for the 90th percentile predictions were higher than those for the measured full-shift exposure data regarding liquids with VP > 10 Pa, powders, and solid objects, whereas opposite results were observed for the other categories (Table 2). The correlations were either high or moderate for all categories except powders; The correlation for powders could not be calculated because the exposure estimate was the same for the two ESs included. The results of bias and relative bias indicated that overall the tool underestimated exposures for all categories except powders. The accuracy for exposures to liquids with VP > 10 Pa was poorer than for solid objects but better compared to the one for the other categories. The value for precision was highest for exposures to liquids with VP > 10 Pa indicating less precise estimates of the tool in this category compared to the others. When individual measurements were compared with the tool estimates, all exposure categories except for powders showed %M>Ts greater than 25% for all percentile estimates (i.e., 50th, 75th, and 90th percentiles).

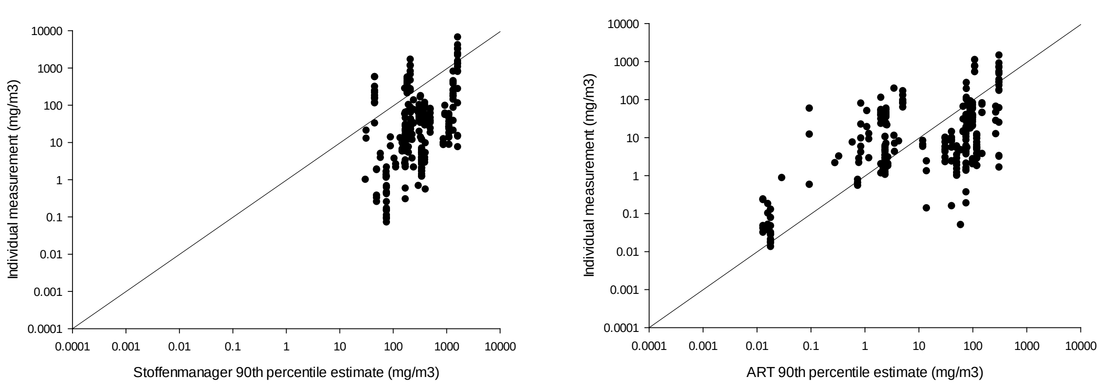

Figure 1. Measured data vs. tools' 90th percentile estimates for exposures to liquids with VP > 10 Pa (Left: Stoffenmanager; Right: ART)

Table 3. Comparison of the exposure measurements with the corresponding tools' exposure estimates (by exposure category)

Tool	Exposure Category	ES No	n	r_p	Bias	Precision	Relative bias (%)	%M>T		
								>50th	>75th	>90th
Stoffenmanager	Aqueous Solutions	2	4	-0.8	1.5	1.3	332	0	0	0
	Liquids with VP ≤ 10 Pa	2	5	0.9 *	-0.8	0.6	-56	100	20	0
	Liquids with VP > 10Pa	42	251	0.4*	0.3	2.1	29	46	22	15
	Powders	2	11	0.1	3.0	1.3	2003	0	0	0
	Solid Objects	3	11	**	**	**	**	NC	NC	NC
ART	Aqueous Solutions	2	4	1.05*	-1.8	0.7	-84	100	100	75
	Liquids with VP ≤ 10 Pa	2	5	0.9*	-1.7	0.5	-82	100	100	80
	Liquids with VP > 10Pa	42	251	0.6*	-0.7	2.3	-53	65	53	41
	Powders	2	11	***	3.8	1.3	4172	0	0	0
	Solid Objects	3	11	0.5	-0.4	1.9	-33	64	45	27

Abbreviation: ES No= number of exposure situations (ESs) for which data were available; n=number of exposure measurements; r_p = Pearson correlation coefficient (log-transformed data); %M>T = Percentage of exposure measurements exceeding the tool estimates; NC = Not calculated because the tool estimates were zero. *p-value < 0. 05; **Stoffenmanager score for handling solid objects was zero and therefore no log-transformation was possible;*** the correlation cannot be calculated because of the same exposure estimate for two ESs.

2) By handling description

Table 4 shows a summary of the performance of higher tier tools by handling/activity description for exposures to liquids with VP > 10 Pa where a number of exposure measurements is greater than 10. Stoffenmanager overall underestimated exposures for the LS (Handling of liquids on large surfaces or large work pieces) task along with a high %M>T (83%) for the 90th percentile estimates. For the other tasks, the positive biases and the comparison of individual measurements with the 90th percentile estimates (ranging from 0 to 15%M>T) suggest that the tool overestimates exposures. The correlation was high for the LPLS (Handling of liquids using low pressure, low speed, or on medium sized surfaces) and either negative or low for the other tasks.

Overall, ART underestimated exposures for all selected tasks except NAF (Activities with relatively undisturbed surfaces [no aerosol formation]), and the results of the comparison between individual measurement and the tool's 90th percentile estimates also indicated and underestimation of exposures (all %M>Ts greater than 30%). A moderate or high correlation was observed for all tasks except SLP (Spreading of liquid products) and NAF (Activities with relatively undisturbed surfaces [no aerosol formation]). The relative bias for NAF was the highest among tasks but note that this was based on only one ES.

Table 4. Summary of the tools' performance by handling/activity description for exposures to liquids with VP > 10 Pa (n > 10)

Tool	Handling/Activity	ES No	n	r_p	Bias	Precision	Relative Bias (%)	%M>T		
								>50th	>75th	>90th
Stoffenmanager	LPHS	4	56	0.2	1.4	1.3	319	16	4	0
	LS	7	23	0.2	-3.0	1.4	-95	96	91	83
	SS	4	30	-0.1	0.8	1.3	116	30	10	0
	LPLS	16	124	087*	0.2	2.0	28	52	22	15
ART	HCO	3	25	0.6*	-0.3	1.8	-26	72	52	36
	FL	12	62	0.9*	-1.4	1.2	-76	92	85	65
	HC&A	7	22	0.4	-3.4	1.4	-97	100	91	86
	SLP	15	109	-0.03	-0.6	2.4	-47	58	40	31
	NAF	1	15	**	2.2	0.6	785	0	0	0

Abbreviation: ES No= number of exposure situations (ESs) for which data were available; n=number of exposure measurements; r_p = Pearson correlation coefficient (log-transformed data); LPHS-Handling of liquids (using low pressure, but high speed) without creating a mist or spray/haze; LS- Handling of liquids on large surfaces or large work pieces; SS- Handling of liquids on small surfaces or incidental handling of liquid; LPLS- Handling of liquids using low pressure, low speed, or on medium sized surfaces; HCO-Handling of contaminated objects or paste, FL-Falling liquids; HC&A-Combined two activities, Handling of contaminated objects or paste & Activities with relatively undisturbed surfaces (no aerosol formation); SLP-Spreading of liquid products; NAF-Activities with relatively undisturbed surfaces (no aerosol formation). *p-value < 0. 05; ** Not calculated because of one ART estimation

3) By input parameters

The impact of input parameters for predicting exposure estimations is summarized in Table 5. Only the results for liquids with VP > 10 Pa are reported. An impact of the LEV input was not calculated for the ART because of the presence of multiple allocations. That is, the majority of ESs (38 out of 51 ESs) has multiple control strategies per ES and could not be simply allocated to either LEV presence or absence.

The %M>T for the Stoffenmanager was zero for the allocation of low VP, while the %M>Ts for the high and medium VP allocation were greater than 15%. The tool's performance was the least accurate for the low VP allocation and the most accurate for the medium VP allocation. Negative correlations were observed for the allocation of medium and low VP. For the LEV input parameter, the tool's accuracy was better when LEV was absent compared to LEV was present, but the observed 18 %M>T in the former allocation indicates that the tool is likely to underestimate exposure levels. None of exposure measurements exceeded the tool's 90th percentile estimates when an LEV was present. The ART's performance for the high VP allocation was the best among the choices of VP input parameter in terms of accuracy. However, the results of negative biases and %M>T greater than 20% for all percentile estimates demonstrated that the tool underestimated exposure levels.

Table 5. Summary of the tools' performance by input parameter for exposures to liquids with VP>10 Pa

Tool	Input parameter		ES No	n	r_p	Bias	Precision	Relative Bias (%)	%M>T		
									>50th	>75th	>90th
Stoffenmanager	Vapour Pressure[(1)]	High	6	37	0.6*	-0.5	1.8	-37	62	38	19
		Medium	31	190	-0.1	0.1	2.0	13	47	20	16
		Low	5	24	-0.4*	2.4	1.8	1046	13	8	0
	LEV	Yes	7	50	0.7*	1.7	1.8	431	24	4	0
		No	35	201	0.2*	-0.1	2.0	-9	51	26	18
ART	Vapour Pressure[(1)]	High	6	37	0.5*	-0.05	1.6	-5	57	35	22
		Medium	31	190	0.1	-0.7	2.4	-51	62	51	39
		Low	5	24	0.9*	-2.1	0.9	-87	100	96	88

Abbreviation: ES No= number of exposure situations (ESs) for which data were available; n=number of exposure measurements; r_p = Pearson correlation coefficient (log-transformed data); [(1)]Low vapour pressure: < 500Pa at room temperature; Medium vapour pressure: 500≤ VP≤10000 Pa; High vapour pressure: VP> 10000 Pa. *Impact of LEV input parameter was not calculated due to multiple control methods per ES for the majority of ESs

4) ESs having only one task

As shown in Table 6only 13 ESs (2 ESs each from the aqueous solutions and liquids with VP ≤ 10 Pa and 9 ESs from the liquids with VP > 10 Pa) of 51 ESs available included a single task,. Because this sib-set of data is rather limited, this section is presented to compare the performance of the Stoffenmanager and ART, and not to draw any conclusion for each tool's performance. Hence, no results of %M/T are reported.

For exposure of aqueous solutions and liquids with VP ≤ 10 Pa, results for both tools were comparable to the ones of the main analysis (Table 3) because the same number of ESs and exposure measurements were considered. For aqueous solutions, Stoffenmanager showed noticeably higher relative bias than ART, whereas for liquids with VP ≤ 10 Pa, an opposite result was observed. For liquids with VP > 10 Pa, both tools showed relative bias < 20%.

Table 6. Comparison of the exposure measurements with the corresponding tools' exposure estimates for ESs having only one task (by exposure category)

Tool	Exposure Category	ES No	n	r_p	Bias	Precision	Relative Bias (%)
Stoffenmanager	Aqueous Solutions	2	4	-0.8	1.5	1.3	332
	Liquids with VP ≤ 10 Pa	2	5	0.9*	-0.8	0.6	-56
	Liquids with VP > 10Pa	9	44	0.5*	0.05	1.3	5
ART	Aqueous Solutions	2	4	1.0*	-1.8	0.7	-84
	Liquids with VP ≤ 10 Pa	2	5	0.9*	-1.7	0.5	-82
	Liquids with VP > 10Pa	9	44	0.3*	0.15	2.2	16

Abbreviation: ES No= number of exposure situations (ESs) for which data were available; n=number of exposure measurements; r_p = Pearson correlation coefficient (log-transformed data); %M>T = Percentage of exposure measurements exceeding the tool estimates. *p-value < 0. 05

DISCUSSION

Description of workplace measurement data

As addressed in Part I (Lee et al., 2017), because most exposure measurements and ESs were available for exposures to liquids with VP > 10 Pa, the imbalanced data for the other exposure categories (n = from 4 to 11 with 2 or 3 ESs) might generate inconclusive results. This might be the case even for exposures to liquids with VP > 10 Pa when performance was determined based on handling or activity descriptions. When the exposure situations were stratified by handling/activity descriptions, only four handling tasks (LPHS, LS, SS, and LPLS) for the Stoffenmanager and five activity tasks (HCO, FL, HC&A, SLP, and NAF) for the ART included more than 10 exposure measurements for use in the comparisons. . Therefore, it is necessary to collect additional exposure measurements for those exposure categories and handling/activity descriptions having limited sample sizes. In addition, i the measured exposure levels for all categories except liquids with VP > 10 Pa were in general low which could be a limitation of the study. Nevertheless, the current study forms one of the most valuable evaluations for the tools because almost all contextual information required for the tool's input parameters was collected during the field surveys. Hence, the findings in this study may be close to the real circumstances in terms of the tool's performance, especially for exposures to liquids with VP > 10 Pa for which sample sizes were large.

Comparison of exposure measurements with tool estimates

<u>Stoffenmanager</u>: A summary of relative bias and %M>T from previous studies and the current study is presented in Table 7. For exposure categories of aqueous solutions and powders, 0%M>T for the tool's percentile estimates (50^{th}, 75^{th}, and 90^{th} percentiles) and positive biases indicated that the tool overestimated exposures levels. The results of previous studies by Koppisch et al. (2012) and van Tongeren et al. (2017) for powders suggest the tool to overestimate exposure, which is in concordance with the findings of the present study. Schinkel et al. (2010) reported 29%M>T for the same category, indicating underestimation of the tool but this result was based on a previous calibration before refining the tool's algorithm. The relative bias of the present study for powders is considerably higher than those from other studies indicating less accurate results. In the present study, exposure measurements for powders were collected from only 2 ESs that might be possible to have biased results. For exposures to liquids with VP ≤ 10 Pa, the tool underestimated exposures (i.e., negative bias) on average but overestimated when the 90^{th} percentile estimates were compared with individual measurements (%M>T = 0%).

Based on these findings, it seemed that the tool's 90th percentile estimates for the aforementioned exposure categories demonstrated promising results, but these results were based on the limited number of exposure measurements (n ≤ 11) from less than or equal to 3 ESs. It is therefore necessary to collect more exposure measurements covering diverse exposure situations to draw firmed conclusions.

For exposures to liquids with VP > 10 Pa having sufficient number of exposure measurements, a positive bias of 0.3 and a relative bias of 29% demonstrated overestimation of the tool on average with an excellent accuracy (Table 3). The calculated precision of 2.1 was similar to those reported by Koppisch et al. (2012) and Schinkel et al. (2010) ranging from 1.5 to 1.8. When individual exposure measurements were compared with the tool's 90th percentile estimates, the proportion of M>T was close to 10%, suggesting the tool to overestimate the 90th percentile of the exposure measurement distribution. Surprisingly, the previous study by van Tongeren et al. (2017) and the present study showed almost identical %M>T for this category. Vink et al. (2010) evaluated the Stoffenmanager v4.0 using 1-methoxypropan-2-ol measurements (n=745 extracted from German MEGA database) collected during professional spraying paint works and reported that the tool's 75th percentile estimates were lower than the exposure measurements for 3 out of 4 tasks (75%). The tool in the present study also underestimated exposures but with lower %M>T (22%). Landberg et al. (2015) tested high and low volatile liquid chemicals from 8 tasks and reported ~27%M>T, whereas the present study showed less percent (%M>T = ~ 14%). Overall, the tool appeared to be accurate and robust enough to predict exposure estimate for this category, and the findings of this study may lead users in favour of adopting this tool.

Table 7. Summary of %M>T and relative bias (%) for the Stoffenmanager

Exposure category	Percentage of exposure measurements exceeding the tool 90th percentile estimates (%M>T)					Relative Bias (%)		
	Ref 1	Ref 2	Ref 3	Ref 4	Present	Ref 1	Ref 2	Present
Aqueous Solutions	N/A	N/A	N/A	N/A	0 (n=4)	N/A	N/A	332
Liquids with VP ≤ 10 Pa	15 (n=40)	N/A	38 (n=8)	26 (n=210)	0 (n=5)	-62	N/A	-56
Liquids with VP > 10Pa	7 (n=72)	N/A	22 (n=27)	14 (n=1326)	15 (n=251)	-11	N/A	29
Powders	29 (n=82)	11 (n=390)	N/A	0.9 (n=118)	0 (n=11)	-77	-25	2003
Solid Objects	N/A	N/A	N/A	N/A	NC	N/A	N/A	NC

Ref 1-Schinkel et al. (2010); Ref 2 – Koppisch et al. (2012); Ref 3 – Spinaze et al. (2017); Ref 4 – van Tongeren et al. (2017); n=number of exposure measurements; N/A = Not applicable; NC = Not calculated because of zero exposure estimation of the tool

For the category of solid objects, workers handled finished solid bars wires for all 3 ESs (ESs 8, 10 and 12) and the tool assumed almost no exposure emission (based on the physical form), leading to zero modelled exposure levels. In practice, we observed exposure ranges from 0.001 to 0.083 mg/m^3 (GM 0.02 mg/m^3) for a full-shift and from 0.002 to 0.33 mg/m^3 (GM 0.03 mg/m^3) for a task-based exposure, although the exposure ranges were low. It seemed that the tool did not account background exposure that could be always present in real workplaces. The findings of this study clearly suggest that the tool's mechanism should be reviewed not to predict zero exposure for handling solid objects.

When considering the impact of handling description, it is recommended to review the algorithm of the LS (handling of liquids on large surfaces or large work pieces) task that resulted in underestimation of exposures (based on a negative bias of -3.0 and 83%M>T). The impact of input parameters for liquids with VP > 10 Pa analysis showed that the allocation of high or medium vapour pressure and the LEV absence resulted in higher %M>Ts compared to the other allocation for each input parameter. Lamb et al. (2015) showed similar %M>Ts for the medium (18%) and low (5%) vapour pressure inputs , whereas the %M>T for the high VP was lower (3%) compared to the present study (19%) (Table 5). Interestingly, the allocation of LEV between the Lamb et al. study and the present study resulted in conficting results (%M>T of LEV presence vs absence- 0% vs. 18% [this study] and 30% vs. 2% [Lamb et al.]). One suggestion might be that the tool developer reviews the assigned scores to the allocation of vapour pressure and LEV inputs.

In summary, the Stoffenmanager seems accurate and robust but still there are areas for improvement of the tool. It needs to be noted that the tool's exposure estimates tend to overestimate the real task-based exposures because we adopted the lowest control method to predict exposure estimate when multiple control methods were present. Hence another suggestion for modifying the tool will be to allow it to account for multiple sub-tasks via its web-version. This could be particularly relevant and useful for small and medium sized enterprises where workers are very likely to perform multiple sub-tasks. Frequently however, clear distinctions between different (sub) tasks are difficult to be made. For example, a batch maker performs a batch-making task by performing several sub-tasks with different control strategies (pouring raw material by opening the top of a batch with a flexible duct located right above the opening, mixing the materials in a completely enclosed system, and placing the product into several containers under partially closed system). One way to resolve this may be to allow the tool to simulate multiple sub-tasks (i.e., based on different control methods).

<u>ART</u>: For aqueous solutions, liquids with VP ≤ 10 Pa, and solid objects, observed negative biases and high %M>Ts (≥ 25%) demonstrated that likely ART underestimates exposures (Table 3). For the powders, although no exposure measurements exceeded the tool's estimates (50^{th}, 75^{th}, and 90^{th} percentiles), the relative bias is considerably higher than those estimated for other categories indicating less accurate results. Again, these findings were based on the limited number of ESs and exposure measurements (ES no = 2 or 3 and n ≤ 11) and hence no firm conclusions can be drawn.

For exposures to liquids with VP > 10 Pa, a negative bias of -0.7 and %M>T above 40% for all percentile estimates suggest that the tool's prediction underestimates exposure. These findings are in agreement with Spinazze et al. (2017) who reported ~ 25% of exposure concentrations exceeding the tool's 90^{th} percentile predictions (n=28; exposure data extracted from previous studies) for organic solvents. However, Hofstetter et al. (2013) reported that the ART 50^{th} percentile estimate was about 2.9 times higher than the exposure measurements of toluene in laboratory-based spray painting tasks which is contradictory to the present results which showed ART's 50^{th} percentile estimates to be considerably lower than the exposure measurements.

As expected from the results of exposures to liquids with VP > 10 Pa, all activities but NAF (Activities with relatively undisturbed surfaces [no aerosol formation]) showed %M>Ts greater than 30%, warranting the tool developer's attention for these activities. For the NAF activity, results are inconclusive because only one ES was considered, however the tool generally overestimated exposures. In addition, the tool's performance based on the allocation of vapour pressure input parameter was determined to be poor, showing negative biases and > 20% of M>T for all choices (high, medium, and low VP).

Based on the findings of this study, it is evident that ART underestimates exposures and tool developers should focus on two potential underlying sources: a) calibration of the ART algorithm and b) reviewing assigned scores representing exposure variability. As described above, the present study cannot determine if the tool's underestimation was from calibration errors due to insufficient exposure measurements per exposure situation. The tool developers might want to recalibrate the ART algorithm by adding more exposure measurements collected from various exposure scenarios. The assigned scores for allocations of each input parameter were made based on literature searches and/or expert judgments, if not available. We strongly suggest that the tool developer pays attention to the available resources (e.g., reports, peer-reviewed papers, etc.) and update the tool, if necessary.

One suggestion as a transition solution is that users adopt upper confidence levels (e.g., upper value of 90% CI of 90^{th} percentile estimate) instead of single estimates because these values include not only exposure variability but also statistical uncertainty. The current study evaluated only exposure estimates using a mechanistic tool based on a number of exposure determinants. Combining this estimation with the tool's internal database or the users' own measurements using a Bayesian statistics can reduce its uncertainty. This would be more practical approach for the tool users.

Comparison of the accuracy between the Stoffenmanager and the ART

As shown in Table 6, the ART is more accurate than the Stoffenmanager for exposures to aqueous solutions, whereas the Stoffenmanager is more accurate than the ART for exposures to liquids with VP ≤ 10 Pa. The performance result for exposures to liquids with VP ≤ 10 Pa were consistent with the findings by Spinazze et al. (2017) but as aforementioned, the present study findings were based on limited data sets. For exposures to liquids with VP > 10 Pa, although Stoffenmanager performed slightly better than ART, both tools appeared to be equally accurate showing relative bias less than or equal to 16%, whereas Spinazze et al. reported that the ART was more accurate that the Stoffenmanager for organic solvents.

CONCLUSIONS

As described in Part I (Lee et al., 2017), sufficient number of sample sizes for the present evaluation was available only for exposures to liquids with VP > 10 Pa. For these exposures, it was determined that ART's accuracy is similar to that of Stoffenmanager. However, the high proportions of exposure measurements exceeding the ART 90^{th} percentile estimates would likely discourage users from adopting this tool for their application. It is strongly recommended that the tool developers perform a comprehensive review of the ART's algorithm including calibration and assigned scores of input parameters. In addition, the Stoffenmanager appeared to better predict the 90^{th} percentile of the exposure distributions compared to ART. Both tools were developed based on the same concept of source-receptor approach by considering near- and far-field regions. However, ART offers a broader range of input parameters compared to the Stoffenmanager even for a single input parameter (e.g., control strategy). Therefore, users who cannot obtain all contextual information required for input parameters would be more beneficial to use the Stoffenmanager instead. However, Stoffenmanager still itself needs be improved, particularly concerning the LS task and its input parameters of vapour pressure and LEV for this category.

For the other exposure categories, the current study was limited to low ranges of exposure levels and insufficient number of exposure measurements from small number of ESs. Given this limitation, the present findings strongly suggest needs for further validation studies by covering a broader range of ESs with large exposure data sets.

In addition, the presence of potential errors caused from various uncertainties, such as in the interpretation of the input parameters, exposure measurement data, and inherent sources cannot be totally excluded from our study. However, the uncertainty from exposure measurements is likely rather negligible because all personal exposure measurements were sampled and analyzed according to well established methodologies (e.g., NIOSH or OSHA methods). Similarly, the obtained consensus agreements that were used to run the tools minimise the uncertainty from the interpretation of the tool input parameters. Yet as expert judgement is still involved in the decision making this process could become a source of error, which could not be determined in the present study. Finally, the inherent uncertainty in each tool itself could be a major source of error.

ACKNOWLEDGMENTS AND FUNDING

The authors are sincerely thankful to xxxxxxxxx for reviewing this manuscript prior to journal submission. The authors also would like to offer special thanks to the employees who participated in our study.

This project was funded internally by the National Institute for Occupational Safety and Health (Project CAN number: 939ZXEY).

DISCLAIMER

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

REFERENCES (will be updated)

ECHA, "Guidance on Information requirements and chemical safety assessment. Chapter R.14 : Occupational exposure assessment," ISBN: 978-92-9495-081-9. Available: https://echa.europa.eu/documents/10162/13632/information_requirements_r14_en.pdf.

Fransman W., J. Cherrie, M. van Tongeren, T. Schneider, M. Tischer, J. Schinkel, H. Marquart, N. Warren, S. Spankie, H. Kromhout and E. Tielemans, "Development of a mechanistic model for the Advanced REACH Tool (ART)," *Ann. Occup. Hyg.,* vol. 55, no. 9, pp. 957-979, 2011.

Hesse S., St. Hahn, K. Schroeder, I. Mangelsdorf, J. Lamb, M. van Tongeren: Evaluation of Tier 1 Exposure Assessment Models under REACH (eteam) Project - Substudy Report on Uncertainty of Tier 1 Models. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin 2015. ISBN: 978-3-88261-159-5

Hofstetter E., J. Spencer, K. Hiteshew, M. Coutu and M. Neally, "Evaluation of recommended REACH exposure modeling tools and near field, far field model in assessing occupational exposure to toluene from spray paint," *Ann. Occup. Hyg.,* vol. 57, no. 2, p. 210–220, 2013.

Koppisch D., J. Schinkel, S. Gabriel, W. Fransman and E. Tielemans, "Use of the MEGA exposure database for the validation of the Stoffenmanager model," *Ann. Occup. Hyg.,* vol. 56, no. 4, pp. 426-39, 2012.

Lamb J., B. G. Miller, L. MacCalman, S. Rashid, M. van Tongeren: Evaluation of Tier 1 Exposure Assessment Models under REACH (eteam) Project - Substudy Report on External Validation Exercise. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin ISBN: 978-3-88261-157-1, Available: http://www.baua.de/en/Publications/Expert-Papers/F2303-D16.html. (2015)

Landberg H.E., P. Berg, L. Andersson, U. Bergendorf, J. Karlsson, H. Westberg and H. Tinnerberg, "Comparison and Evaluation of Multiple Users' Usage of the Exposure and Risk Tool: Stoffenmanager 5.1.," *Ann. Occup. Hyg.,* vol. 59, no. 7, pp. 821-835, 2015.

Eun Gyung (Emily) Lee, Judith Lamb, Nenad Savic, Bojan Gasic, Christian Jung, Michael L. Kashon, Jongwoon Kim, Martin Tischer, Martie van Tongeren, David Vernez, Martin Harper, Evaluation of exposure assessment tools under REACH: Part I – Tier 1 models, Annals of Occupational Hygiene, In press, 2017

McDonell P., J. Schinkel, M. Coggins, W. Fransman, H. Kromhout, J. Cherrie and E. Tielemans, "Validation of the dust algorithm of the Advanced REACH Tool using a dataset from the pharmaceutical industry," *Journal of Environmental Monitoring,* vol. 13, no. 6, pp. 1597-1606, 2011.

McNally K., N. Warren, W. Fransman, R. Entink, J. Schinkel, M. van Tongeren, J. Cherrie, H. Kromhout, T. Schneider and E. Tielemans, "Advanced Reach Tool: A Bayesian model of Occupational Exposure Assessment," *Ann. Occup. Hyg. ,* vol. 58, no. 5, pp. 551-565, 2014.

Marquart H., Heussen, H, M. Le Feber, D. Noy, E. Tielemans, J. Schinkel, J. West and D. Van der Schaaf, "'Stoffenmanager', a web-based control banding tool using an exposure process model," *Ann. Occup. Hyg.,* vol. 52, no. 6, pp. 429-441, 2008.

Schinkel J., W. Fransman, H. Heussen, H. Kromhout, H. Marquart and E. Tielemans, "Cross-validation and refinement of the Stoffenmanager as a first tier exposure assessment tool for REACH," *Occup. Environ. Med.,* vol. 67, no. 2, pp. 125-32, 2010.

Schinkel J., N. Warren, W. Fransman, M. van Tongeren, P. McDonnell, E. Voogd, J. Cherrie, M. Tischer, H. Kromhout and E. Tielemans, "Advanced REACH Tool (ART): Calibration of the mechanistic model," *Journal of Environmental Monitoring ,* vol.13, no. 5, pp. 1374-1382, 2011.

Schinkel J., P. Ritchie, H. Goede, W. Fransman, M. van Tongeren, J. Cherrie, E. Tielemans, H. Kromhout and N. Warren, "The Advanced REACH Tool (ART): Incorporation of an Exposure Measurement Database," *Ann. Occup. Hyg.,* vol. 57, no. 6, pp. 717-727, 2013.

Schinkel J., W. Fransman, P. McDonnell, R. Entink, E. Tielemans and H. Kromhout, "Reliability of the Advanced REACH Tool (ART)," *Ann. Occup. Hyg.,* vol. 58, no. 4, pp. 450-468, 2014.

Spinazzè A., F. Lunghini, D., Campagnolo, Sabrina Rovelli, Monica Locatelli, Andrea Cattaneo and Domenico M. Cavallo, "Accuracy evaluation of three modelling tools for occupational exposure assessment," Annals of Work Exposures and Health, 2017 advanced article.

Tielemans E., D. Noy, J. Schinkel, H. Heussen, D. van der Schaaf, J. West and W. Fransman, "Stoffenmanager exposure model: development of a quantitative algorithm," *Ann. Occup. Hyg.,* vol. 52, no. 6, pp. 443-454, 2008.

Tielemans E., N. Warren, W. Fransman, M. van Tongeren, K. McNally, M. Tischer, H. Kromhout, J. Schinkel, T. Schneider and J. Cherrie, "Advanced REACH Tool (ART): overview and research needs," *Ann. Occup. Hyg.,* vol. 55, no. 9, pp. 949- 956, 2011.

Vink S. , J. Mikkers, T. Bouwman, H. Marquart and E. Kroese, "Use of read-across and tiered exposure assessment in risk assessment under REACH – A case study on a phase-in substance," *Regulatory Toxicology and Pharmacology,* vol. 58, no. 1, pp. 64-71, 2010.