Annals of Work Exposures and Health, 2019, Vol. 63, No. 6, 637–650
doi: 10.1093/annweh/wxz027
Advance Access publication 16 May 2019
Original Article

Original Article

Calibration of the Dermal Advanced REACH Tool (dART) Mechanistic Model

Kevin McNally,[1,*] Jean-Philippe Gorce,[1] Henk A. Goede,[2] Jody Schinkel[2] and Nick Warren[1]

[1]HSE Science and Research Centre, Health and Safety Executive, Buxton SK17 9JN, UK; [2]Netherlands Organisation for Applied Scientific Research (TNO), Risk Assessment for Products in Development (RAPID), PO Box 360, 3700 AJ Zeist, The Netherlands

*Author to whom correspondence should be addressed. Tel: +44-0203-028-2066; e-mail: kevin.mcnally@hse.gov.uk

Submitted 31 October 2018; revised 11 March 2019; editorial decision 14 March 2019; revised version accepted 29 April 2019.

Abstract

The dermal Advanced REACH Tool (dART) is a Tier 2 exposure modelling tool currently in development for estimating dermal exposure to the hands (mg min^{-1}) for non-volatile liquid and solids-in-liquid products. The dART builds upon the existing ART framework and describes three mass transport processes [deposition (D_{hands}), direct emission and direct contact (E_{hands}), and contact transfer (T_{hands})] that may each contribute to dermal exposure. The mechanistic model that underpins the dART and its applicability domain has already been described in previous work. This paper describes the process of calibrating the mechanistic model such that the dimensionless score that results from encoding contextual information about a task into the determinants of the dART can be converted into a prediction of exposure (mg min^{-1}). Furthermore, as a consequence of calibration, the uncertainty in a dART prediction may be quantified via a confidence interval. Thirty-six experimental studies were identified that satisfied the conditions of: (i) high-quality contextual information that was sufficient to confidently code the dART mechanistic model determinants; (ii) reliable exposure measurement data sets were available. From these studies, 40 exposure scenarios were subsequently developed. A non-linear log-normal mixed-effect model was fitted to the data set of D_{hands}, E_{hands}, and T_{hands} scores and corresponding measurement data. The dART model was shown to be consistent with activities covering a broad range of tasks [spray applications, activities involving open liquid surfaces (e.g. dipping, mixing), handling of contaminated objects, spreading of liquid products, and transfer of products (e.g. pouring of liquid)]. Exposures resulting from a particular task were each dominated by one or two of the identified mass transport processes. As a consequence of calibration, an estimate of the uncertainty associated with a mechanistic model estimate is available. A 90% multiplicative interval is approximately a factor of six. This represents poorer overall precision than the (inhalation) ART model for dusts and vapours, although better than the ART model for mists. Considering the complexity of the conceptual model compared with the ART, the wide variety of exposure scenarios considered with differing dominant routes, and the particular challenges that

result from the consideration of measurement data both above and beneath a protective glove, the precision of the calibrated dART mechanistic model is reasonable for well-documented exposure scenarios coded by experts. However, as the inputs to the model are based upon user judgement, in practical use, the reliability of predictions will be dependent upon both the competence of users and the quality of contextual information available on an exposure scenario.

Keywords: calibration; dermal exposure; exposure modelling; skin

Introduction

Legislation on the Registration, Evaluation, Authorisation, and Restriction of Chemicals (REACH) entered into force within the member states of the European Union (EU) on 1 June 2007. The legislation makes manufacturers and importers of chemicals responsible for assessing and managing the risk posed by chemicals. In addition to the registration process, a chemical safety assessment must be performed when supply of a substance reaches 10 tonnes per year. This assessment include an exposure assessment and risk characterization if the substance fulfils the criteria for classification according to various hazard classes or categories set out in Annex I to the CLP Regulation [Regulation (EC) No 1272/2008 on the classification, labelling, and packaging of substances and mixtures] or is assessed to be persistent, bioaccumulative, and toxic (PBT) or very persistent and very toxic (vPvB). Due to the very large number of chemicals covered by the legislation, phased deadlines for registration were introduced, with 1 June 2018 being the final deadline.

Many thousands of exposure scenarios need to be assessed under REACH, and exposure modelling provides an essential support to this process (Tielemans *et al.*, 2011; McNally *et al.*, 2014). A pragmatic tiered approach to risk assessment has been proposed for both occupational and environmental settings (OECD, 1997). In this conceptual framework, screening tools are classed as Tier 1 approaches and cover broad classes of exposure scenarios. Through conservative assessments, they are designed to efficiently screen those broad classes of exposure scenarios of no concern from those in need of more detailed (and typically, expensive) assessment. Progressively more complex and specific exposure assessments are performed for substances where the (assumed) conservative estimates from Tier 1 approaches exceed measures of the potential hazard of a substance, quantified by a metric such as a derived no-effect level (DNEL). Higher tiered methodologies progressively incur greater cost and resource requirements compared with lower tiered approaches.

A variety of tools are available for assessing inhalation exposures, with Tier 1 tools recently reviewed by Riedmann *et al.* (2015). The Advanced REACH Tool (ART) (Tielemans *et al.*, 2011; McNally *et al.*, 2014) is a web-based generic Tier 2 tool (www.advancedreachtool.com) that is suitable for assessing exposure scenarios that are not filtered by Tier 1 screening tools. The ART follows a Bayesian approach, making use of mechanistically modelled estimates of exposure for a range of substance classifications, information on exposure variability from meta-analyses in the literature, and any available exposure measurements.

Goede *et al.* (2019) noted that important developments in dermal exposure assessment have taken place over the past two decades. Of particular note are the quantitative models that have been developed for biocides (BEAT) (TNsG, 2007) and industrial chemicals (RISKOFDERM) (Warren *et al.*, 2006). Whilst these tools can be considered as higher tier and in principle can be utilized under REACH, as data-driven models, these tools are not always transparent on their applicability domains (Goede *et al.*, 2019). A small number of Tier 1 screening tools such as ECETOC TRA (ECETOC, 2012) and the Stoffenmanager (Marquart *et al.*, 2008) are also available for assessing dermal exposures. To date, however, no Tier 2 tools are currently available for dermal exposure assessment.

The dART is a new generic Tier 2 exposure model currently in development for estimating dermal exposures to the hands to low volatile liquid products and solids in liquid. The tool is based upon the existing ART platform and incorporates elements of the ART model for estimating dermal contamination by aerosol deposition (one of three routes in the dermal exposure model). The mechanistic model framework is based upon that of the ART (Fransman *et al.*, 2011) and follows a source-receptor model (Schneider *et al.*, 1999) with principal modifying factors (MFs) along the source-receptor pathway. A detailed description of the dART mechanistic model is provided in Goede *et al.* (2019) and the references therein. The focus of the current work is calibration of the dART mechanistic model, which is the process of converting the dimensionless score estimated by the dART mechanistic model, into an exposure estimate following a similar process to that described for the ART in Schinkel *et al.* (2011). The calibrated

mechanistic model of dART provides an estimate of the geometric mean (GM) exposure and is one aspect of the overarching statistical model that underpins the dART. A manuscript describing the full Bayesian model of the dART, which incorporates within- and between-worker sources of variability (thus allowing estimates of percentiles of the exposure distribution) and which can refine estimates based upon dermal exposure measurements, is in preparation.

Methods

dART Conceptual model

The dART conceptual model is described in detail in Goede *et al.* (2019) and the supporting references. Briefly, the model is built on the categorization of workplace activities into six main activity classes (AC): spray applications (AC1), activities involving open liquid surfaces (e.g. dipping, mixing) (AC2), handling of contaminated objects (AC3), spreading of liquid products (AC4), application of liquids in high-speed processes (e.g. oil drilling) (AC5), and transfer of products (e.g. pouring of liquid) (AC6). There is a further level of abstraction for a subset of activity classes; for example, AC1 activities can be subcategorized as surface spraying (AC1.1) and space spraying (AC1.2).

Dermal exposure to the hands is modelled considering five mass transfer processes, originally described by Schneider *et al.* (1999). These five processes include aerosol deposition from the air (D_{hands}), direct emission from splashes (E_{hands}), transfer through hand-to-surface contacts (T_{hands}), retention capacity of the hands (or gloved hands), and protection provided by chemical protective gloves. For each of the three exposure routes (D_{hands}, E_{hands}, and T_{hands}), the conceptual model defines independent principal MFs that underlay the exposure mechanism. An MF may be itself composed of several underlying multiplying determinants. Each determinant is assigned a score within a range based upon expert judgement with assessments made relative to a baseline category for the determinant. For example, the weight fraction of the substance has eight levels in the conceptual model (starting from 'Pure substance' (100%) down to 'Minute' (0.01–0.1%)) with respective scores ranging from 1 (for pure substance) down to 0.0006 (for minute amount). For this determinant, 'Pure substance' represents the baseline level (i.e. score of 1).

Once multiplied together, the MFs generate an overall score for each route of exposure. Due to the multiplicative method used to derive the scores for each route of exposure, the scores inevitably cover several orders of magnitude (up to seven orders of magnitude). In some cases, the model predicts zero exposure from a particular route. The scores for each route are dimensionless and quantify the relative exposures in any set of exposure scenarios. A calibration model, described in the section Calibration model, is required to convert these dimensionless scores to an estimate of exposure to the hands in terms of mass of product accumulated over the duration of activity.

Data sets, assignment of scores, and reliability

Thirty-six distinct experimental studies investigating hand exposure to low volatile liquids were identified. Field studies were selected from the Bayesian Exposure Assessment Toolkit [BEAT, Biocides Steering Group (2002)] and Bystanders, Residents, Operators and Workers Exposure models for plant protection products (BROWSE, www.browseproject.eu) databases, which contain measurement data, summary information, and images from high-quality dermal exposure experimental studies. The original field studies were reviewed for scoring when necessary (for BROWSE scenarios, the detailed information recorded for each measurement was sufficient to code the determinants). Summary detail on the work tasks, sampling methods, and references to the experimental studies are given in the online Supplementary Material.

The 36 experimental studies selected were used to refine and calibrate the scoring systems developed by Goede *et al.* (2019). These studies contained sufficient contextual information of the workplaces and work processes taking place (with photographs) to assign scores to each MF and underlying multipliers confidently. The descriptions of the sampling methodologies and analytical methods were judged adequate and the hand exposure measurements were reported with a sufficient level of detail (i.e. raw measurement data including sampling duration, dilution of analyte in the applied product, and mass of analyte were available for each individual) to be used for the dART model calibration.

Scoring a scenario

Information required for the appropriate scoring of the scenario included if the activity took place indoors or outdoors, in open space or next to buildings or a densely planted area. For indoor exposure scenarios, information on both the room volume and ventilation characteristics were required. Information on the product use, including the amount of product applied, its concentration, and viscosity the pressure at which the product may be applied, the surface area treated, duration, and orientation of application were required. Finally, information on the level of segregation, distance from source (arm length, handheld tool), the level of automation,

frequency of contact with tools, controls, objects treated or secondary surfaces encountered during the task, and complexity of the objects being treated were required.

Scores for each MF were assigned according to the scheme described in Goede *et al.* (2019). A score, for D_{hands}, T_{hands}, and E_{hands} was calculated on a per measurement basis based upon the contextual information and pictures from the 36 experimental studies. In the majority of scenarios, the scores were common to all the individual measurements available within a unique exposure scenario. However, within-scenario scores varied by up to a factor of 10 from the median. This small within-scenario variation in scores resulted from a single determinant taking different values (e.g. some workers may have used a short handle brush while others used a longer brush). Relatively homogenous scores were desirable for calibration since significant variations in the scores assigned to measurements would imply that the measurements did not correspond to the same exposure scenario.

The scores for all scenarios were coded by a single member of the team (JG) with approximately 25% of scenarios, including at least one scenario per activity class, checked by the primary mechanistic model developer (HG) to ensure that the contextual information associated with tasks was being correctly transcribed into the determinants of the dART mechanistic model.

Refinement of mechanistic model

An iterative process was adopted in model development with each iteration comprising of the steps: (i) development or refinement of the mechanistic model; (ii) calibration of the non-dimensional dART score using representative exposure measurements; and (iii) evaluation of model fit. Refinement of the model was driven by the evaluation stage. After each calibration, the model output was examined: evidence of lack-of-fit was assessed for and provided evidence for refinement of the mechanistic model. The focus of this analysis was not on particular scenarios where the model under- or over-predicted exposures, although severe lack-of-fit of individual scenarios was examined, but rather on activity classes where the model appeared to be in error.

Whilst information obtained from model calibration was used to establish deficiencies in the mechanistic model and provided guidance on whether the mechanistic model had a tendency to under- or over-predict, the changes to the model were informed by the literature. Additional research was conducted in order to establish an evidence base for changes to the mechanistic model. It is important to note that the scores assigned to the categories of the various MFs in the mechanistic model were not simply tuned such that measurements associated with the scenarios were consistent with predictions from the calibrated mechanistic model. As an example of this process of refinement, one significant improvement to the model resulted after preliminary calibration results suggested exposures for a subset of spraying scenarios, where the exposed individual was separated from the source by an enclosed tractor cab, were several orders of magnitude in excess of model predictions. Additional data sets were identified, which corroborated this finding, and a literature search was conducted to inform the required revisions to the mechanistic model—a consideration of secondary sources of surface contamination (Ramwell *et al.*, 2005). The possibility of encountering multiple contaminated secondary surfaces was subsequently considered in all scenarios. The coding in this work is based on the final version of the mechanistic model at the conclusion of model development.

Data extraction, manipulation, and version control

The calculations required to convert the scored determinants into scores for D_{hands}, T_{hands} and E_{hands} were conducted using Microsoft Excel spreadsheets. New spreadsheets in date-stamped subdirectories were created for each new version of the mechanistic model. After the completion of scoring, the spreadsheets were saved to .csv files and read in to the R software (The R Foundation for Statistical Computing, 2014). All manipulations and data cleaning were performed using R scripts: a direct audit trail between primary data and the final analysis data set was therefore documented.

A modelling data set was constructed by extracting the relevant information from these input files. The three dART dimensionless scores for D_{hands}, T_{hands}, and E_{hands} were extracted into the modelling data set. Information on each experimental measurement relating to the exposure scenario was also extracted. The primary variable extracted was of the measurement of the active substance on the hands divided by the sampling time, yielding an exposure 'to mass of active substance' per minute of activity (mg min^{-1}). Further information on the sampling method (cotton sampling gloves, hand washes, etc.), and personal protective equipment (PPE) worn, the activity duration, and the concentration of the measured substance in the applied product were also extracted.

All experimental measurements (to active substance) were normalized by the weight fraction of the active substance present in the applied product used during the activity. The primary variable in statistical analysis was of 'exposure to mass of product' per minute of activity (mg min^{-1}). A second derived variable, volume of product per minute of exposure (µl min^{-1}, derived as mass per minute

of exposure divided by density of product) was also calculated. Loadings onto the hands or gloved hands (exposure per minute multiplied by sampling duration) in terms of both mass and volume of product were finally calculated.

Some summary information on the 40 exposure scenarios coded from the 36 experimental studies is provided in Table 1: more detailed descriptions of the studies are provided in the online Supplementary Material. For most scenarios the scores for D_{hands}, T_{hands}, and E_{hands} were common to all measurements within the scenario; in the scenarios where there were differences in scores between measurements the tabulated values represent median scores for determinants. Two sets of summary statistics are also provided for the associated measurements; the GM and geometric standard deviation (GSD) (mg min^{-1}) for measurements 'not under PPE', which includes measurements taken above a glove (potential exposures), using a protective glove as the sampling device (potential exposures) or on bare hands in the absence of a protective glove (actual exposures), and 'under PPE' (actual exposures). For most scenarios, one column is blank which reflects that sampling tended to be either exclusively 'under PPE' or 'not under PPE' within a scenario. In only four scenarios (scenario labels 1, 5, 29, and 30) were measurements above and beneath a protective glove available. The AC 1.1 scenario 'Knapsack motorized spraying' (scenario 5) was unique in that matched-pair measurements (a measurement both above and beneath a protective glove) for 15 workers were available.

A comparison of the exposure to product (mg min^{-1}) in the 40 coded exposure scenarios is made in Fig. 1. The boxplots in the upper panel of Fig. 1 are based upon experimental measurements taken 'not under PPE', whereas the boxplots in the lower panel of Fig. 1 represent measurements taken 'under PPE'. In scenarios 1, 5, 29, and 30, measurements under the protective glove were approximately two orders of magnitude lower than those taken above the protective glove.

The exploratory analysis of the derived modelling data set revealed two data sets where the measurements were unreliable. The first data set, AC4 scenario 'car-washing' (data shown under scenario 32 in Fig. 1), involved the repeated saturation of the bare hands. The measurement data suggested that the mass of analyte on the hands (measured by a hand wash) was independent of sampling duration and there was very little variability in the measurements. An unpublished data set from HSE Buxton estimates a maximum retention of 4 ml (approximately 4 g) on the skin resulting from dipping water-like substances: the relatively low measurements in the scenario were assessed to be as a consequence of the low retention capacity of the bare skin. The calculated rate of exposure to product per minute was therefore a potentially large underestimate. The 12 measurements in this scenario were therefore treated as right-censored data (i.e. the true exposure was greater than that experimentally measured). The second data set, AC2.2 scenario 'Electroplating' (Scenario 21 in Table 1 and Fig. 1), used portable X-ray fluorescence spectrometry (PXRF) as the analytical method (Roff *et al.*, 2004). Only one of the 22 glove measurements in this study was above the relatively high limit of detection from this method. The measurements in this scenario were therefore treated as left-censored data (i.e. the true exposure was less than the limit of detection).

Calibration model

Calibration of the ART model for inhalation exposures is described in Schinkel *et al.* (2011). A conceptually similar process is followed here although there is the intermediate step of deriving a single dART score from the three exposure routes. The deposition, transfer, and direct emission routes have an associated score for every exposure scenario. The scores for each route are dimensionless and quantify the relative exposures in any set of exposure scenarios but not absolute exposure; i.e. a score for scenario A of twice that in scenario B estimates that the exposure due to a particular route in scenario A is twice that of the exposure scenario B. However, the length scales of the three scores and the overall importance of the routes differ. A re-scaling of the transfer and emission scores, via constants β_1 and β_2, respectively, was made such that the importance of these routes was assessed relative to deposition. These constants are estimated within the calibration and account for differences in the scales for the respective scores and the relative importance of the routes in determining the dermal exposure.

A single dART score for every exposure scenario is derived from equation (1), where, as noted above, parameters β_1 and β_2 are estimated during calibration. The importance of each route for a given exposure scenario is given by the fractions of the dART score corresponding to D_{hands}, $\beta_1 T_{hands}$ and $\beta_2 E_{hands}$, respectively.

$$\text{DART score} = D_{hands} + \beta_1 T_{hands} + \beta_2 E_{hands} \quad (1)$$

The conceptual model that converts a relative exposure quantified by a dART score to an estimate of mass of product per minute of exposure is achieved through calibration of the dART scores using exposure measurements. The model described in Schinkel *et al.* (2011) and adapted in present work is of the form of equation (2). The additional scaling parameter α, estimated through the calibration model, converts the

Table 1. A comparison of median dART scores for the 40 exposure scenarios and summary statistics on the measurements (mg min^{-1}).

Class	Label	Description	*N*	D_{BP}	T_{BP}	E_{BP}	GM (GSD)	
							Not under PPE	Under PPE
AC1.1	1	Anti-fouling spraying	25	2.27250	1.51500	1.35000	62.1 (1.9)	0.50 (6.9)
AC1.1	2	Car body spraying	29	0.04848	0.00150	0.45000	2.33 (2.6)	NA
AC1.1	3	Anti-fouling paint spraying	12	12.1200	0.00450	0.45000	41.2 (3.6)	NA
AC1.1	4	Spray cleaning foam	12	2.70810	0.03150	4.50000	36.1 (2.9)	NA
AC1.1	5	Knapsack motorized spraying	30	7.57500	0.00315	0.45000	28.1 (2.2)	0.14 (2.7)
AC1.1	6	Knapsack spraying	9	0.20200	0.00315	0.05000	8.95 (6.2)	NA
AC1.1	7	Spraying low-pressure lance	6	11.1100	0.11445	0.45000	NA	1.35 (2.8)
AC1.1	8	Spraying high-pressure lance	26	99.9900	0.19950	4.05000	33.6 (2.6)	NA
AC1.1	9	Insecticide spraying	20	11.1100	0.09450	4.50000	72.6 (5.8)	NA
AC1.1	10	Boom spraying closed cabin	6	0.00152	0.03535	0.00000	1.42 (1.8)	NA
AC1.1	11	Boom spraying semi-closed cabin	11	0.04545	0.03850	0.00000	10.2 (4.8)	NA
AC1.1	12	Boom spraying closed cabin	4	0.00152	0.03535	0.00000	1.03 (1.9)	NA
AC1.1	13	Boom spraying semi-closed cabin	7	0.00354	0.03535	0.00000	1.20 (4.2)	NA
AC1.1	14	Boom spraying closed cabin	7	0.10605	0.03850	0.00000	5.83 (3.5)	NA
AC1.1	15	Boom spraying semi-closed cabin	6	0.00354	0.03535	0.00000	5.11 (5.7)	NA
AC1.1	16	Boom spraying in orchard cabin	7	0.01515	0.03850	0.00000	1.65 (4.5)	NA
AC1.1	17	Boom spraying in orchard no cabin	2	0.15150	0.03850	0.00000	13.2 (1.1)	NA
AC1.1	18	Boom spraying in orchard cabin	4	0.00101	0.03535	0.00000	0.36 (3.2)	NA
AC1.2	19	Fogging	8	26.9333	0.00315	0.00500	NA	0.03 (5.3)
AC2.1	20	Dipping activities (timber)	4	0.01239	10.5000	500.000	NA	5.54 (4.2)
AC2.2	21	Electroplating	22	0.26933	0.10500	4.50000	NA	1.16 (1.3)
AC2.2	22	Electroplating (KRIOH)	29	0.26933	0.10500	4.50000	NA	0.17 (1.8)
AC2.2	23	Electroplating (KRIOH)	4	0.26933	0.01050	1.50000	NA	0.34 (1.1)
AC3	24	Timber pre-treatment (solvent)	19	0.00081	1.05000	1.35000	NA	0.28 (8.8)
AC3	25	Timber pre-treatment (water)	50	0.00269	10.5000	40.5000	NA	2.86 (5.2)
AC3	26	AF net deployment (solvent)	5	2.02E-05	1.35000	3.64500	NA	0.14 (1.5)
AC3	27	AF net deployment (water)	3	6.73E-05	0.32319	12.1500	NA	0.33 (1.5)
AC3	28	Forestry: packing and planting	9	0.00060	0.31500	0.45000	NA	0.01 (6.2)
AC4	29	Non-professional brush painting	13	0.22725	0.00150	0.45000	6.0 (6.0)	0.3 (11.5)
AC4	30	Brush and roller painting	10	0.22725	0.19500	12.1500	73 (1.1)	0.8 (10.1)
AC4	31	Opening and closing packaging, cleaning equipment and spreading	30	0.27270	0.00150	0.45000	1.0 (10.5)	NA
AC4	32	Car washing	12	0.80800	95.5500	250.000	399 (1.5)	NA
AC4	33	Large-scale surface wiping	24	2.72700	13.5000	250.000	1908 (1.9)	NA
AC6.2	34	Loading DEGBE	28	0.00727	0.02000	13.5000	45 (16.7)	NA
AC6.2	35	Filling of spray guns	28	0.00394	0.02000	1.50000	21.5 (4.6)	NA
AC6.2	36	Filling DEGBE	9	0.02182	0.02000	4.50000	68.1 (3.5)	NA
AC6.2	37	Filling DEGBE	10	0.02424	0.00583	0.45000	1.08 (7.8)	NA
AC6.2	38	Filling DEGBE	1	0.00242	0.00515	0.00000	33.0 (NA)	NA
AC6.2	39	Filling DEGBE	4	0.02424	0.00650	0.45000	4.59 (3.0)	NA
AC6.2	40	Filling DEGBE	6	0.08080	0.01400	5.00000	17.7 (4.4)	NA

relative score to an exposure estimate in the units of the measurements (mg min^{-1}).

$$\text{Exposure} = \alpha \times \text{DART score} \quad (2)$$

As in Schinkel *et al.* (2011), the statistical model relating experimental measurements to dART scores is specified on the log scale, which reflects an assumption that variability in measurements is proportional to magnitude

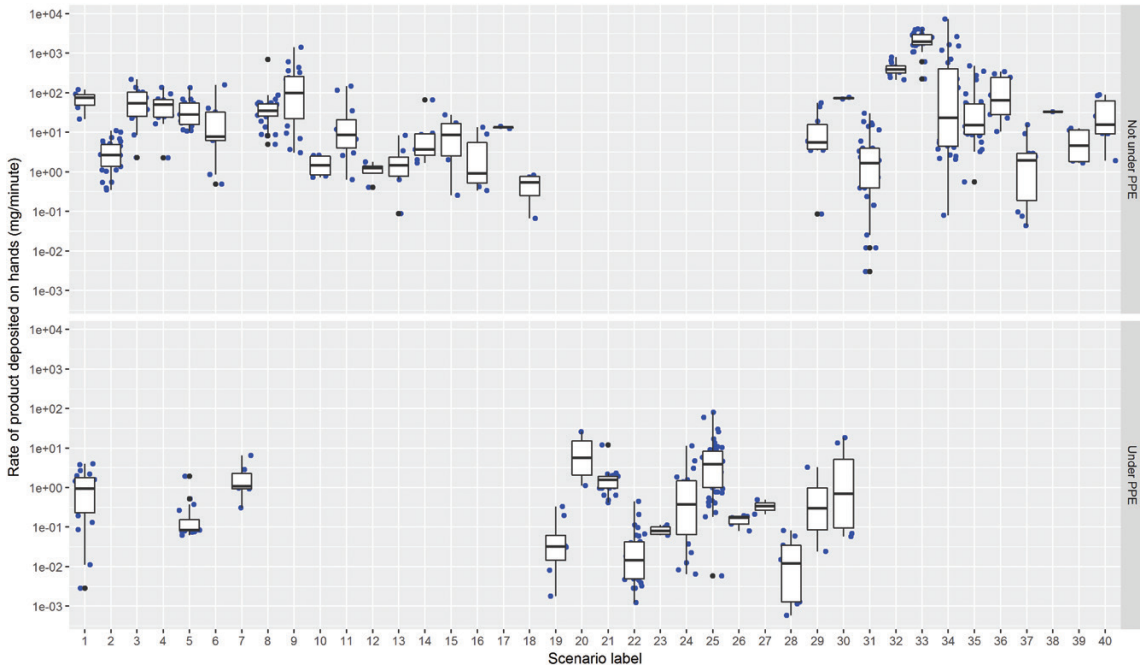

Figure 1. A comparison of the rate of product deposited onto the hands (mg min^{-1}) in the 40 coded exposure scenarios. Boxplots represent summaries of the measurements, the points represent individual measurements.

(McNally *et al.*, 2014). A non-linear mixed-effect model, shown in equation (3), was used in current work.

$$\log\left(Y_{ij}\right) = \log(\alpha) + \log\left(D_{\text{hands}} + \beta_1 T_{\text{hands}} + \beta_2 E_{\text{hands}}\right) + I_G \times \text{Glove} + s_i + \varepsilon_{ij} \quad (3)$$

$$s_i \sim N(0, \sigma_s)$$

$$\varepsilon_{ij} \sim N(0, \sigma_r)$$

In the model, shown in equation (3), Y_{ij} denotes the j^{th} measurement in the i^{th} scenario. The fixed effect parameters α, β_1, and β_2 are described above. The final fixed-effect parameter Glove and the associated indicator variable I_G account for the potential protection offered by a protective glove (typically a nitrile glove). I_G was coded as one for measurements sampling beneath a PPE glove and zero otherwise (which included cases where a sampling glove was worn above a PPE glove, the PPE glove was used as the sampling device, and a PPE glove was not worn). On the log scale, the protection offered due to glove is additive; therefore, on the natural scale, the efficacy of the glove is multiplicative [with factor exp (Glove)].

The final two terms in equation (3) account for systematic and residual variability, respectively. The mechanistic model is an imperfect approximation to reality and as a consequence will underestimate or overestimate exposures for a given set of determinants D_{hands}, T_{hands}, and E_{hands}. A random-effect s_i estimated for each scenario (40 additional parameters in all) quantifies the average (systematic) mechanistic model error for each of the scenarios. Term ε_{ij} represents the j^{th} residual error in the i^{th} scenario and includes between-worker and within-worker sources of variability: consistent information was not available across all scenarios to allow the more precise modelling of components of variance considered in Schinkel *et al.* (2011). Random effects and residual errors were assumed independent and normally distributed with standard deviations of σ_s and σ_r, respectively.

The separation of variability in measurements into residual and systematic variability is the correct mathematical treatment. However, the quantification of systematic error is significant for the development of a tool for predicting dermal exposures. The random-effect standard deviation σ_s summarizes the systematic error in the mechanistic model predictions for the calibration scenarios. As long as the calibration scenarios and the associated errors (under- or over-predictions) are a representative subset of exposure scenarios, and there are no unexplained trends, the potential error in a mechanistic model prediction in scenarios where measurements are unavailable can be quantified. This allows an estimate of dermal exposure

and probability bounds to be provided. Refinements to the mechanistic model made during each iteration of the model development process aimed to reduce the systematic error. Systematic trends in random effects by activity class, i.e. consistent underestimation or overestimation by the mechanistic model for a particular class or classes, were used to inform refinement of the mechanistic model.

During the model development phase, the models were fitted using the nlme package of R (The R Foundation for Statistical Computing, 2014) with parameters estimated using restricted maximum likelihood estimation (REML). During this phase, the censoring in scenarios 21 and 32, as described in section Data extraction, manipulation, and version control, was not accounted for—the measurements were treated as observations. The final calibration model was fitted within a Bayesian framework using the Winbugs software (Lunn *et al.*, 2009) and the censored observations in scenarios 21 and 32 were correctly modelled as left and right-censored measurements, respectively. Gaussian priors were used for parametersα, β_1, and β_2 and Glovewith means of zero and standard deviations of 10 000 (such that the priors were only weakly informative); however, β_1 and $\beta_2$were constrained to be positive (in part to ensure numerical stability since negative values for these parameters could result in the expression within the log taking a negative value; however, this assumption also models a belief that the three exposure routes each contribute to dermal exposure and negative coefficients do not reflect this belief). Non-informative gamma prior distributions with shape and scale parameters set to 0.01 were assumed for $\sigma_s$andσ_r. Inference was made using a Markov Chain Monte Carlo (MCMC) algorithm. The chain was burnt in for 10 000 iterations and sampling was conducted for a further 10 000 iterations. Every 10th sample was retained for inference.

An approach based upon Bayesian utility theory was adopted in order to select a single, optimal set of model parameters—optimal in the sense that the parameter set provided a good fit to measurements and was unbiased for each activity class (i.e. the tendency to consistently underestimate or overestimate for any activity class was minimized). The loss function, shown in equation (4), was computed for each of the retained parameter sets in the MCMC sample. In equation (4), μ_i represents the mean of the random effects for the i^{th} activity class, n_i the number of scenarios in the i^{th} activity class, and σ_s is the between-scenario standard deviation [equation (3)]. A separate μ_i was computed for the boom and surface spraying scenarios (AC 1.1). The summation in equation (4) is over the activity classes.

$$L = \exp(-0.5 \sum n_i \sigma_s^{-2} (\mu_i - 0)^2) \quad (4)$$

The posterior density, the loss function, and the posterior loss (loss weighted by the probability density) were calculated for the 1000 retained parameter sets described above with the parameter set that minimized posterior loss chosen as the optimal parameter set.

Variant models

The analysis described above was repeated for the alternate dependent variable of volume of product per minute of activity (µl min^{-1}). The majority of scenarios involved a water-like product and assumed a density of 1 g cm^{-3} (1 mg translates to 1 µl). However, for four scenarios, the applied product had a density significantly greater than 1 g cm^{-3}: the product in use occupied a smaller volume compared with a water-like substance.

A second alternative model that encoded a belief that the protective effect of PPE gloves was in proportion to the dermal challenge, rather than as a multiplicative factor, was also considered. Under this model, PPE gloves offer a proportionately greater protective effect for scenarios resulting in a very large dermal exposure. Formally this model is written as equation (5) with random effect and residual error terms as before.

$$\log(Y_{ij}) = \log(\alpha) + (1 - I_G \times \text{Glove}) \times \log(D_{\text{hands}} + \beta_1 T_{\text{hands}} + \beta_2 E_{\text{hands}}) + s_i + \varepsilon_{ij} \quad (5)$$

The modelling approach detailed above was repeated for both of these alternative models and differences compared with the primary model are reported in results.

Results

Summary statistics (posterior median and a 90% credible interval) for the parameters in the calibration model calculated from the MCMC output are given in Table 2; note these are summaries of the respective marginal posterior distributions—the model corresponding to all

Table 2. Posterior median and a 90% credible interval for parameters in the initial calibration model.

Parameter	Median	90% credible interval
α	−1.74	−3.21, −0.24
β_1	522	113, 1779
β_2	29.6	5.3, 172.5
Glove	−5.3	−5.8, −4.8
σ_s	1.24	0.98, 1.60
σ_r	1.61	1.53, 1.69

parameters set at their respective posterior medians does not necessarily correspond to the 'best-fitting' parameter set (nor indeed necessarily to a plausible parameter set). The parameter values corresponding to the optimal model (that which minimized posterior loss) are given in Table 3.

The three key fixed-effect terms that determined the relative importance of the three routes were all highly uncertain. Pairwise correlations calculated from the simulations were –0.85, –0.77, and 0.72 for α andβ_1,α andβ_2, and$\beta_1$and β_1, respectively. The credible interval for Glove was relatively narrow; however, on the natural scale this range corresponds to a 90% credible interval of a 120- to 330-fold reduction in exposures when sampling under chemical protective gloves. The credible intervals for the between-scenario and residual error standard deviations were relatively narrow. Residual error (variation) was greater than systematic (between-scenario) error.

A point estimate of the GM exposure (mg min^{-1}) for the 40 calibration scenarios summarized in Table 1 was calculated using the calibration model, written as equation (3) with optimized coefficients from Table 3. These estimates of exposure and the fractions of exposure due to the deposition, direct emission, and transfer routes are given in Table 4. Note that for this comparison, which examines the general behaviour of the calibrated model, the effect of a chemical protective glove is not considered. The results (Table 4) indicated that the scores for particular activity classes were driven by one or two of the routes. Notably, the deposition route was only important for the spraying scenarios (AC1.1 and AC1.2).

A comparison of estimates of exposure for each of the calibration scenarios with the GM calculated from the corresponding data set (Table 1) is made in Fig. 2. In this comparison, the effect of a protective glove is taken into account where appropriate, based upon the coefficient for Glove in Table 3. A point estimate from the model and a 95% credible interval based on the random-effect standard deviation of 1.08 (Table 3) is shown alongside each of the predictions made in Fig. 2. Note that for these predictions, the central estimates are based upon contextual information about the scenario encoded as three scores (Table 1). The standard deviation is based upon typical scenario errors resulting from application of the conceptual model, written as equation (3), and demonstrates practical use. This treatment clearly differs from a mixed-effect model based upon measurements alone, where the magnitude of error would be scenario-specific and typically much smaller.

Table 3. Optimized model parameters—mass per minute of exposure.

(log(α)	β_1	β_2	*Glove*	σ_s	σ_r
0.13	87.40	5.02	–5.72	1.08	1.68

For two of the scenarios where measurements were available above/without a protective glove, the GM of the measurements was outside the credible interval. The calibrated model underestimated the exposure for the scenario 'knapsack spraying', with the central estimate under-predicting the GM by a factor of 10. The model overestimated the GM for scenario 'car-washing'; however, for this scenario, the hands were regarded as being saturated and all measurements were treated as right censored in the analysis. A considerable underestimation of the GM is therefore reasonable and in line with how this scenario was coded in the statistical model.

The performance of the model for measurements underneath a chemical protective glove was poorer, with the GM lying above the credible interval for four scenarios and beneath the credible interval for one scenario. In all these cases, the GM was relatively close to the bounds of the credible interval. The greatest under-prediction corresponded to the scenario 'Electroplating', where 26 of the 27 measurements were coded as left censored and therefore an under-prediction is reasonable. Poorer performance of the model applied to scenarios where sampling was beneath PPE is not surprising because the use of a protective glove is an additional complication. The degree of protection offered by protective gloves will vary with a number of factors Creely and Cherrie (2001) and the large protective factor of a 300-fold reduction in exposure estimated in the calibration model could be a considerable overestimate for some situations. There was also evidence from the calculated GSDs (and an examination of the raw data for these scenarios) (Table 1) that contamination of the sampling gloves (probably during removal of contaminated chemical protective gloves) could have occurred for some measurements, thus inflating the GM—this observation is based on the measurement made on a single glove of a single worker per scenario being a significant outlier compared with other measurement data.

Variant models

The analysis was repeated for the alternate dependent variable of volume of product per minute of activity (µl min^{-1}). Only a brief summary of results is presented for this model.

There was a modest increase in the importance of deposition in the model for volume compared with the model for mass. However, the direct emission

Table 4. Estimated exposures (mg min^{-1}) from the optimized calibration model for the 40 exposure scenarios and fractions of exposure due to the deposition, transfer, and direct emission routes.

Class	Description	D_{BP}	T_{BP}	E_{BP}	Estimate (mg min^{-1})
AC1.1	Anti-fouling spraying	0.016	0.936	0.048	161.92
AC1.1	Car body spraying	0.022	0.054	0.928	2.79
AC1.1	Anti-fouling paint spraying	0.820	0.027	0.153	16.9
AC1.1	Spray cleaning foam	0.097	0.098	0.805	32.11
AC1.1	Knapsack motorized spraying	0.748	0.028	0.224	11.57
AC1.1	Knapsack spraying	0.277	0.386	0.349	0.83
AC1.1	Spraying low-pressure lance	0.475	0.428	0.097	26.74
AC1.1	Spraying high-pressure lance	0.726	0.127	0.148	157.61
AC1.1	Insecticide spraying	0.265	0.197	0.539	48.02
AC1.1	Boom spraying closed cabin	0	1	0	3.54
AC1.1	Boom spraying semi-closed cabin	0.013	0.987	0	3.9
AC1.1	Boom spraying closed cabin	0	1	0	3.54
AC1.1	Boom spraying semi-closed cabin	0	1	0	3.54
AC1.1	Boom spraying closed cabin	0.030	0.97	0	3.97
AC1.1	Boom spraying semi-closed cabin	0	1	0	3.54
AC1.1	Boom spraying in orchard cabin	0.005	0.995	0	3.87
AC1.1	Boom spraying in orchard no cabin	0.042	0.958	0	4.02
AC1.1	Boom spraying in orchard cabin	0	1	0	3.54
AC1.2	Fogging	0.989	0.01	0.001	31.15
AC2.1	Dipping activities (timber)	0	0.268	0.732	3923.8
AC2.2	Electroplating	0.008	0.286	0.705	36.67
AC2.2	Electroplating (KRIOH)	0.008	0.286	0.705	36.67
AC2.2	Electroplating (KRIOH)	0.031	0.105	0.864	9.98
AC3	Timber pre-treatment (solvent)	0	0.931	0.069	112.8
AC3	Timber pre-treatment (water)	0	0.819	0.181	1283.18
AC3	AF net deployment (solvent)	0	0.866	0.134	156
AC3	AF net deployment (water)	0	0.316	0.684	102.15
AC3	Forestry: packing and planting	0	0.924	0.076	34.1
AC4	Non-professional brush painting	0.087	0.050	0.863	3
AC4	Brush and roller painting	0.003	0.218	0.779	89.59
AC4	Opening and closing packaging, cleaning equipment, and spreading	0.102	0.049	0.849	3.05
AC4	Car washing	0	0.869	0.131	10 996.53
AC4	Large-scale surface wiping	0.001	0.484	0.515	2790.35
AC6.2	Loading DEGBE	0	0.025	0.975	79.59
AC6.2	Filling of spray guns	0	0.188	0.811	10.63
AC6.2	Filling DEGBE	0.001	0.072	0.927	27.89
AC6.2	Filling DEGBE	0.009	0.181	0.809	3.2
AC6.2	Filling DEGBE	0	1	0	0.52
AC6.2	Filling DEGBE	0.009	0.199	0.794	3.26
AC6.2	Filling DEGBE	0.003	0.046	0.950	30.23

component was notably smaller for the calibration based upon volume—the scores for this route were approximately a half of those based upon calibration to mass. For the modelling data set, this difference had the greatest impact on the scenarios involving transfer of liquids (AC6.2) where the direct emssion route was consistenly the most influential; however, for AC4 scenarios involving hand immersion, this could have an important effect on predictions. Fig. 3 shows a comparison of exposure estimates per minute of activity for calibrations based upon mass and volume. The comparison indicates that the differences between models based on mass and those based on volume were generally small. At most, the estimates could differ by a

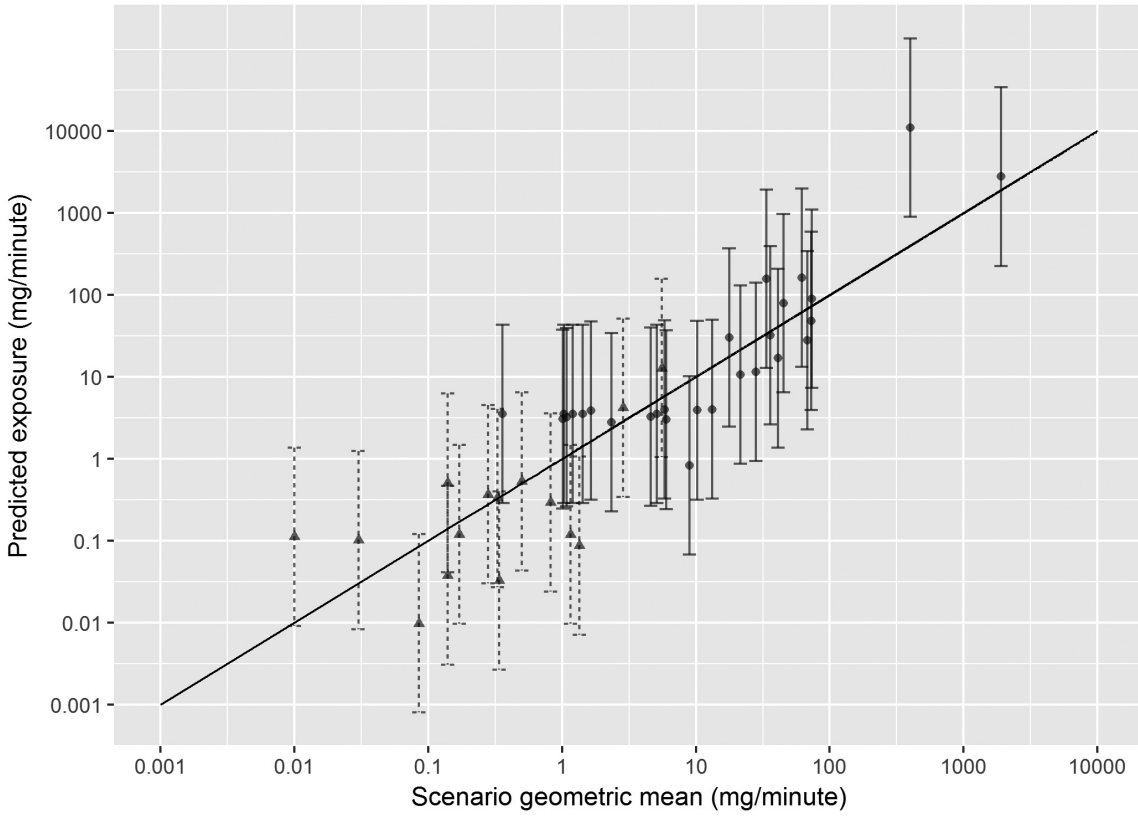

Figure 2. A comparison of the central estimate and a 95% credible interval for mass of product per minute of activity with the GM calculated from measurements: beneath PPE (triangles and dashed lines); above/no PPE (circles and solid lines). One- to-one line indicating perfect correspondence is also shown.

factor of two (for scenarios dominated by direct emission), a difference which is small compared with the overall range of predictions.

The alternative modelling of the effect of a protective glove [equation (5)] resulted in a notably poorer fit to the data.

Discussion

In this work, a calibration model has been developed for estimating a rate of dermal exposure to the hands to low volatile liquid products and solids in liquid based upon three scores, D_{hands}, T_{hands}, and E_{hands}, defined in the mechanistic model of the dART (Goede *et al.*, 2019). Calibration models for mass of product (mg min^{-1}) and volume of product (µl min^{-1}) as the dependent variable in statistical models were both considered. A similar quality of fit to the data was achieved for both models, principally since the density of the product in use was close to 1 g cm^{-3} for the vast majority of exposure scenarios identified. Broadly similar estimates of the calibrated dART scores were obtained (Fig. 3). Whilst both models for mass and volume of product can be used, the recommendation from this work is that predictions of mass of product per minute of exposure to the hands are made using the dART on the basis that the analysis of variance components conducted by Kromhout *et al.* (2004), which feeds into the overarching statistical model of the dART, was also based upon mass (mg min^{-1}). However, for a given scenario, a conversion from a dART prediction of mg min^{-1} to a prediction of volume can be easily made, given the density of the applied product.

Based upon the model written as equation (3) and calibrated parameters (Table 3), a central estimate of the GM with probability bounds can be estimated for dermal exposure to the hands (per minute of exposure) for tasks involving low volatility liquids and solids in liquid. The between-scenario standard deviation estimated in the calibration models is a critical part of this work since it allows the uncertainty in the central estimate of a dART prediction to be quantified. For the suggested model of mass of product per minute of exposure,

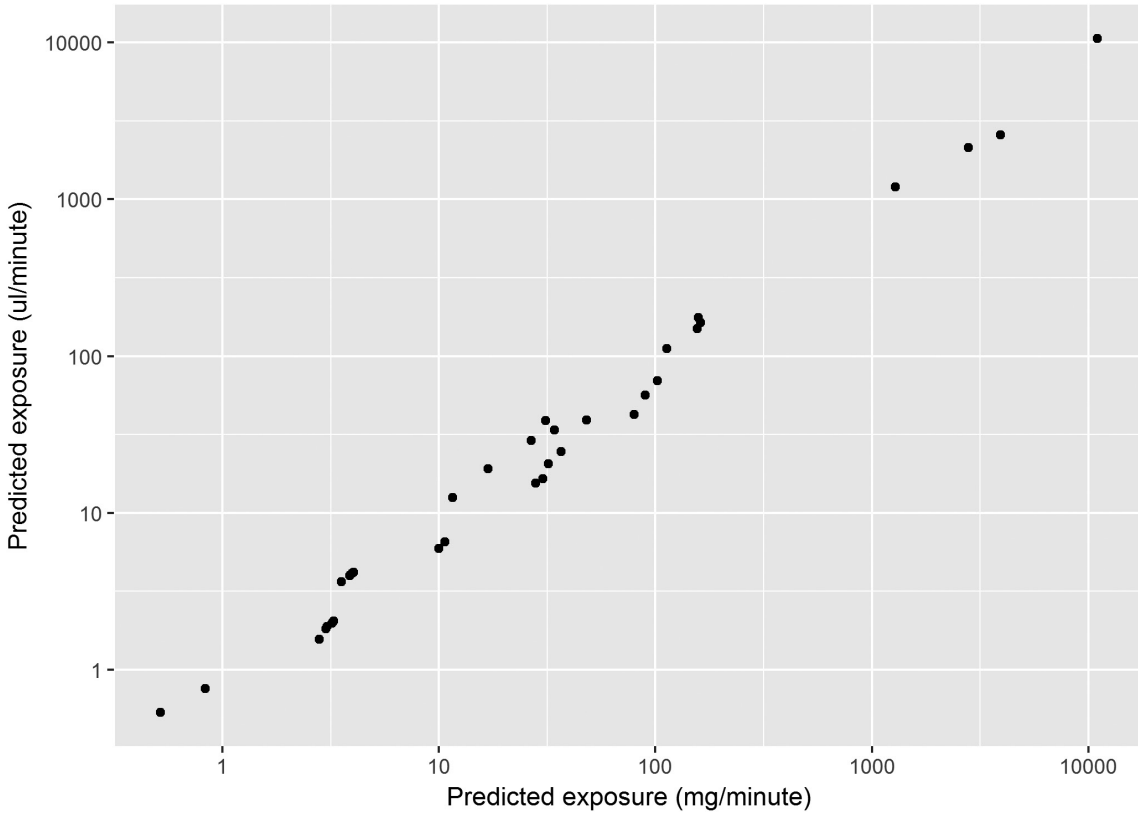

Figure 3. A comparison of central estimates of exposure based upon calibrated models of mass and volume per minute of exposure.

a 90% multiplicative interval is approximately a factor of six. This represents poorer overall precision than the ART model for dusts and vapours, although better than the ART model for mists. Considering the complexity of the conceptual model compared with the ART, this represents a good outcome.

Whilst the estimate of exposure from these calibrated models is in terms of total product per minute of activity, an estimate of the active ingredient is easily obtained by multiplying the estimated product loading rate by the proportion of active ingredient in the product in use. Predictions would be of exposure to the ingredient or mixture of ingredients of concern.

It was necessary to include a statistical parameter to account for the use of chemical protective gloves since the calibration data set comprised of measurements both above protective gloves (or based on sampling of bare hands) and beneath protective gloves. For models formulated for both mass and volume of product on the hands, the estimates for the variable Glove suggested that measurements beneath a protective glove were a factor of 300 below those taken either above of or in the absence of a protective glove.

Although the chemical barrier resistance of protective gloves to chemicals is tested under laboratory conditions (British Standards Institution, 2016), such a test is only relevant for the selection of the right chemical protective glove material in relation to the chemical used but does not take into consideration other parameters relevant to the control of exposure. Those parameters may include inspecting gloves for physical damage before using them, exposing the gloves to mixtures of chemicals, changing gloves regularly once used (i.e. typically at least every 8 h or at the end of every work shift), donning and doffing gloves in a safe manner and selecting the right gloves for the job (e.g. right size, no excessive loss of dexterity, right cuff length). For that reason, tests on protective gloves in such control conditions only provide a theoretical optimum protection level that might not be offered to the wearer of such gloves in workplace settings. The actual level of protection is difficult to evaluate prior to the use of the gloves. Investigations on the protection levels offered by PPE gloves suggest that human

behaviour may be a dominant factor when considering the control of exposure to workers' hands (Garrod *et al.*, 2001; Rawson *et al.*, 2005). Some assessments of the effectiveness of protective gloves calculated from the ratio of outer and inner samplers have been reported in the literature. Creely and Cherrie (2001) calculated protection factors ranging from 96 to 470 (or equivalently a reduction in exposure of between 98.9 and 99.8%) relative to bare skin.

Whilst the 300-fold reduction for measurements taken beneath a protective glove (estimated by the statistical model) is consistent with the experimental work of Creely and Cherrie (2001), an interpretation of this statistical parameter as an indicator of the efficacy of protective gloves is overly simplistic. The calibrated dART model represents an estimate of product deposited onto the hands and does not currently take account of removal from the skin [the fourth of the mass transport processes in the conceptual model of Schneider *et al.* (1999)]. The retention of product onto the sampling device may differ from that of the skin (or a chemical protective glove). For the experimental studies identified for this work, cotton sampling gloves were used in the vast majority of cases and thus measurements are likely to represent an overestimate of exposure to the hand (Gorman Ng, 2014). Whilst this has no implications for the calibration of the dART model, some account needs to be taken of retention in practical use. The conceptual model described in Goede *et al.* (2019) considers this determinant.

The mechanistic model developed for the dART is complex, reflecting the three identified routes of dermal exposure and with the model and scoring routine having gone through a number of iterations. Detailed information was required to parameterize the determinants; however, it was difficult to obtain high-quality studies where all the necessary information was available. Even in the detailed studies selected for this work, a degree of expert judgement was required to choose the appropriate score for some determinants. A further difficulty arose from the possible interaction of workers with the sampling device (typically a cotton sampling glove). Workers can unintentionally contaminate a sampling device during donning and doffing. In some scenarios, it was not clear whether protective gloves had been removed at some point during the task creating opportunities for undue contamination of the sampler. The inflated GSD for measurements taken under a protective in four scenarios (labels 1, 24, 29, 30) were as of a consequence of a single large measurement corresponding to a single hand from an individual worker (note that the large GSD in over-glove measurements for scenarios 31 and 34 corresponded to tasks where the dART scores within the scenario differed: the large variability in measurements in these scenarios reflects this).

Despite the small number of exposure scenarios used in the calibration, these did span a breadth of classes (Table 1) and an analysis of the relative importance of the D_{hands}, T_{hands}, and E_{hands} scores in the overall dART score (Table 4) suggests the results are in line with expectations. However, there is clearly a need for further testing of the model using data sets that have not been utilized in the model development process and potentially a refinement of the parameter estimates should suitable exposure scenarios become available. Ideally, scenarios for the activity classes AC5 and AC6.1 that were not utilized in calibration would be identified. Blind testing of the model on scenarios that have not been used in the calibration would be a powerful test of the predictive ability of the model too.

Calibration requires high-quality contextual data on a task so that scoring can be made with confidence and reliable D_{hands}, T_{hands}, and E_{hands} scores compared with measurements on the task through the calibration model. However, in practical use, conservative assumptions (resulting in a higher dART score) can be made by a user of the tool when contextual information related to some determinants is weak. Such an approach ensures that where there is uncertainty, it translates to an overestimate rather than an underestimate of dermal exposure. Whilst this suggestion is a practical approach to accommodating imperfect information, it would be impractical when the entire task is poorly documented.

The dART is a task-specific model for dermal exposure to the hands. The calibration process focussed on exposure scenarios where task-specific measurement data sets were available and hence the measured dermal exposure could be attributed to a specific task. In principle, more complex data sets could have been used so long as exposure periods associated with tasks were available and some account was taken of potential loss of product when transitioning between tasks or periods of non-exposure. Estimates of exposure associated with a job, comprising multiple tasks leading to dermal exposure, can be made by taking a time-weighted sum of task-specific exposure estimates. A period of no exposure can also be incorporated into this approach.

Considering the complexity of the conceptual model compared with the ART, the wide variety of exposure scenarios considered with differing dominant routes and the particular challenges that result from the sampling device not being independent of the worker, the precision of the calibrated dART mechanistic model is reasonable

when expert users are coding well-documented exposure scenarios. However, as the inputs to the model are based upon user judgement, in practical use, the reliability of predictions will be dependent upon both the competence of users and the quality of contextual information available on an exposure scenario. Validation of the model using verification data sets not used in model calibration is recommended once the statistical framework of the dART model has been finalized.

Supplementary Data

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

Funding

This research was funded by the Health and Safety Executive and the Dutch Ministry of Social Affairs and Employment. The paper's contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

Conflict of interest

The authors declare no conflict of interest.

References

Biocides Steering Group (2002). Final report: guidance on exposure assessment of biocides. Available at https://echa.europa.eu/documents/10162/16960215/bpd_guid_tnsg+human+exposure+2002_en.pdf. Accessed 30 October 2018.

British Standards Institution (2016). BS EN 374-1:2016 Protective gloves against dangerous chemicals and micro-organisms. *Terminology and performance requirements for chemical risks*. Hemel Hempstead, UK: BSI Standards Publication.

Creely KS, Cherrie JW. (2001) A novel method of assessing the effectiveness of protective gloves–results from a pilot study. *Ann Occup Hyg*; **45**: 137–43.

ECETOC. (2012) *European Centre for Ecotoxicology and Toxicology of Chemicals. ECETOC TRA version 3: background and rationale for the improvements*. Technical report, No. 114. Brussels, Belgium: ECETOC.

Fransman W, Van Tongeren M, Cherrie JW *et al*. (2011). Advanced REACH tool (ART): development of the mechanistic model. *Ann Occup Hyg*; **55**: 957–79.

Garrod AN, Phillips AM, Pemberton JA. (2001) Potential exposure of hands inside protective gloves-a summary of data from non-agricultural pesticide surveys. *Ann Occup Hyg*; **45**: 55–60.

Goede HA, McNally K, Gorce JP *et al*. (2019). Dermal advanced REACH tool (dART) – development of a hand exposure model for low-volatile liquid products. *Ann Work Expo Health*; **63**: 624–636.

Gorman Ng M, de Poot S, Schmid K *et al*. (2014) A preliminary comparison of three dermal exposure sampling methods: rinses, wipes and cotton gloves. *Environ Sci Processes Impacts*; **16**:141–7.

Kromhout H, Fransman W, Vermeulen R *et al*. (2004). Variability of task-based dermal exposure measurements from a variety of workplaces. *Ann Occup Hyg*; **48**: 187–96.

Lunn D, Spiegelhalter D, Thomas A *et al*. (2009). The BUGS project: evolution, critique and future directions. *Stat Med*; **28**: 3049–67.

Marquart H, Heussen H, Le Feber M *et al*. (2008). 'Stoffenmanger', a web-based control banding tool using an exposure process model. *Ann Occup Hyg*; **52**:429–41.

McNally K, Warren N, Fransman W *et al*. (2014). Advanced REACH tool: a bayesian model for occupational exposure assessment. *Ann Occup Hyg*; **58**: 551–65.

OECD. (1997). Guidance document for the conduct of studies of occupational exposure to pesticides during agricultural application. OECD series on testing and assessment No. 9. Available at http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ocde/gd(97)148&doclanguage=en. Accessed 30 October 2018.

Ramwell CT, Johnson PD, Boxall ABA *et al*. (2005). Pesticide residues of the external surfaces of field crop sprayers: occupational exposure. *Ann Occup Hyg*; **49**: 345–50.

Rawson BV, Cocker J, Evans PG *et al*. (2005). Internal contamination of gloves: routes and consequences. *Ann Occup Hyg*; **49**: 535–41.

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.

Roff M, Bagon DA, Chambers H *et al*. (2004). Dermal exposure to dry powder spray paints using PXRF and the method of Dirichlet tesselations. *Ann Occup Hyg*; **48**: 257–65.

Schinkel J, Warren N, Fransman W *et al*. (2011) Advanced REACH tool (ART): calibration of the mechanistic model. *J Environ Monit*; **13**: 1374–82.

Schneider T, Vermeulen R, Brouwer DH *et al*. (1999). Conceptual model for assessment of dermal exposure. *Occup Environ Med*; **56**: 765–73.

The R Foundation for Statistical Computing. (2014). R Version 3.0.3 (2014-03-06).

Tielemans E, Warren N, Fransman W *et al*. (2011) Advanced REACH tool (ART): overview of version 1.0 and research needs. *Ann Occup Hyg*; **55**: 949–56

TNsG. (2007). *Human exposure to biocidal products. Technical notes for guidance*. Helsinki, Finland: ECHA. Available at https://echa.europa.eu/documents/10162/16960215/bpd_guid_tnsg-human-exposure-2007_en.pdf. Accessed 30 October 2018.

Warren N, Marquart H, Christopher Y *et al*. (2006) Task-based dermal exposure models for regulatory risk assessment. *Ann Occup Hyg*; **50**: 491–503