Regulatory Toxicology and Pharmacology 121 (2021) 104864

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Compounded conservatism in European re-entry worker risk assessment of pesticides

Felix M. Kluxen [a,*], Edgars Felkers [a], Jenny Baumann [b], Neil Morgan [c], Christiane Wiemann [d], Franz Stauber [e], Christian Strupp [f], Sarah Adham [g], Christian J. Kuster [b]

[a] ADAMA Deutschland GmbH, Cologne, Germany
[b] Bayer AG, Crop Science Division, Monheim, Germany
[c] Syngenta, Bracknell, United Kingdom
[d] BASF Oesterreich GmbH, Agricultural Solutions, Vienna, Austria
[e] BASF SE, Agricultural Solutions, Ludwigshafen, Germany
[f] Gowan Crop Protection, Reading, United Kingdom
[g] Corteva Agriscience, Abingdon, United Kingdom

ARTICLE INFO

Handling editor: Dr. Martin Van den berg

Keywords:
Risk assessment
Re-entry
Worker
Compounded conservatism
Default values
Pesticides
Plant protection products, Non-dietary exposure

ABSTRACT

We review the risk parameters and drivers in the current European Union (EU) worker risk assessment for pesticides, for example considering crop maintenance, crop inspection or harvesting activities, and show that the current approach is very conservative due to multiple worst-case default assumptions.

As a case study, we compare generic exposure model estimates with measured worker re-entry exposure values which shows that external cumulative exposure is overpredicted by about 50-fold on average. For this exercise, data from 16 good laboratory practice (GLP)-compliant worker exposure studies in 6 crops were evaluated with a total number of 184 workers.

As generic overprediction does not allow efficient risk management or realistic risk communication, we investigate how external exposure can be better predicted within the generic model, and outline options for possible improvements in the current methodology. We show that simply using averages achieves more meaningful exposure estimates, while still being conservative, with an average exposure overprediction of about 9-fold.

Overall, EU risk assessment includes several numerically unaccounted "hidden safety factors", which means that workers are well protected; but simultaneously risk assessments are biased towards failing due to compounded conservatism. This should be considered for further global or regional guidance developments and performing more exposure-relevant risk assessment.

1. Introduction

Regulatory agencies worldwide determine whether pesticides can be safely used. In the European Union (EU), one key aspect of Regulation (EC) 1107/2009 concerning placing plant protection products (PPP) on the market is to ensure a high level of protection for humans.

For all relevant populations that may be exposed to either the pesticide itself or residues and degradation products thereof, risk assessments are mandatory to demonstrate a lack of concern. To achieve this goal a hazard-based reference dose is set to put into context the exposure dose which is usually estimated from generic predictive models.

To that end, it is necessary to consider a safety margin of at least 100 when comparing the exposure with the reference doses. In practice, an experimentally derived hazard characterization dose value is divided by a safety assessment or uncertainty factor (AF/UF) of at least 100 to generate reference values that are intended to account for both interspecies and intraspecies differences. It should be noted that these doses are typically considered to be "no adverse effect levels" and depending on the dose spacing in toxicological studies, this may add more conservatism into the setting of reference values. When such reference values are derived from toxicological animal bioassays, the safety margin cannot be reduced according to Regulation (EC) 1107/2009 even if human data demonstrate similar sensitivity towards the effects

* Corresponding author. ADAMA Deutschland GmbH, Edmund-Rumpler Str. 6, 51149, Köln, Cologne, Germany.
E-mail address: Felix.Kluxen@adama.com (F.M. Kluxen).

https://doi.org/10.1016/j.yrtph.2021.104864
Received 20 October 2020; Received in revised form 18 December 2020; Accepted 8 January 2021
Available online 12 January 2021

observed in test animals. Furthermore, toxicological studies carried out on humans should not be used to lower the safety margins for active ingredients (AI) or plant protection products (Reg. 1107/2009 (13) and Article 8 §2). Other regulations allow UFs below 100, e.g. the European Chemicals "REACh" Regulation (1907/2006 on Registration, Evaluation, Authorisation and Restriction of Chemicals) or pesticide regulations from other geographical regions, e.g. the United States of America, which also allow the use of human data or pharmacokinetic/dynamic (PKPD) modelling to evaluate necessary UFs.

Regulation (EC) 1107/2009 does not elucidate what the intended protection goal is and how a "high level of protection" for humans would be otherwise characterized. Thus, an UF of 100 can be assumed to correspond to an about 100 nominal margin of exposure (towards the adverse effect levels), which seems to be considered sufficiently protective to classify a product as "safe" within the European Community.

The no observed adverse effect level (NOAEL) is used as a surrogate adverse effect level, conversely, the benchmark dose (BMD), relates to a certain benchmark response (BMR), which is considered biologically relevant or adverse. If a NOAEL cannot be derived in an appropriate toxicological study, also the lowest observed adverse effect level (LOAEL) can be used by considering additional safety factors. It should be noted that the effect level concept has been criticized for being dependent on the actually tested doses and because NOAELs are derived by statistical methods which are subject to power considerations (Haber et al., 2018; Hardy et al., 2017; Jensen et al., 2019; Kluxen, 2020).

Some UFs used in human risk assessment are based on toxicokinetic considerations (Dankovic et al., 2015). However, the common UF of 100, which is used within Regulation (EC) 1107/2009, was generically set (possibly based on Lehman and Fitzhugh (1954)) and only subsequently reinterpreted to relate to toxicokinetic/toxicodynamic inter- and intra-species differences (Dourson and Stara, 1983; WHO, 1994; WHO, 2005). UFs are intended to cover both the uncertainty associated with deriving the hazard characteristic estimate and biological variation between species used for hazard assessment and also variability in the human target population (Dankovic et al., 2015; Hayes, 2014; Ragas, 2011; Vermeire et al., 1999). Depending on the severity of effects or the steepness of the dose-response observed in the toxicological studies, or other uncertainties defined in the hazard data package, additional (e.g. 3X) arbitrary safety factors may be used in deriving reference values.

Further uncertainty factors are also included when characterizing the exposure with modelling approaches. In Europe, uncertainty associated with non-dietary exposure estimates is additionally accounted for by using high empirical percentiles or other statistical approaches (EFSA, 2014; EFSA, 2017b).

Since uncertainty is taken into account both in hazard and exposure assessment, it is reasonable to assume that the "intended" margin of exposure of at least 100 is substantially exceeded by compounding in the European non-dietary risk assessment approach. To derive realistic risk estimates and allow efficient risk management and appropriate risk communication, it is thus helpful to shed light on the conservatisms in the current EU risk assessment approach for plant protection products. In this paper, this is primarily achieved by investigating the means by which exposure is estimated and the associated default assumptions. This exercise provides greater transparency and may offer authorities options to better refine risk management and puts the results of non-dietary risk assessments into a more relevant and meaningful perspective. We quantify the level of conservatism later with the example of external dose estimation for worker re-entry exposure and compare model predictions with measured exposure data from several worker exposure studies.

Cochran and Ross (2017) previously discussed in detail how uncertainties in pesticide risk assessment are enumerated, leading to much higher actual margins of safety than 100. In order to mention some aspects of hidden conservatism in pesticide risk assessments that are not quantified, they discussed conservatism associated with deriving toxicological reference values, e.g. due to route-to-route extrapolation and dose spacing or differences in exposure duration between the exposure scenario and the corresponding toxicological studies, and also conservatism in the exposure assessment process. These observations, in conjunction with the precautionary nature of exposure estimation, which we address, illustrate the overall high level of conservatism evident in European risk assessment.

1.1. Worker exposure as a case study

Workers entering areas that were previously treated with plant protection products have been recently associated with an increased risk, since the European Food Safety Authority (EFSA) guidance on non-dietary exposure (2014) and the EFSA Guidance on dermal absorption (EFSA DA guidance) (EFSA, 2012a; EFSA, 2017b) came into force. Before, re-entry worker risk assessments were often uncritical. This may hint at two potential issues in the re-entry risk assessment evaluation in Europe: (i) hazard and exposure assessment guidance documents are developed independently from each other and (ii) theoretical assumptions in the exposure estimation may drive the risk assessment without necessarily identifying an actual new or increased risk. These two suppositions are explored in detail in the manuscript.

While EFSA (2014) uses a regression-based approach on measured exposure data for operator exposure, re-entry exposure is estimated by multiplying generic default values, which is essentially similar to the method of Krebs et al. (2000).

According to EFSA (2014), re-entry exposure is a function of

- the applied dose of active ingredient (a.i.) per hectare (AppRate) in kg a.i. applied/ha,
- potentially a multiplication factor (MAF) that considers multiple similar and successive applications and foliar residue decline,
- the extent of foliar contact during re-entry depending on a specific re-entry scenario - transfer coefficient (TC), in cm^2/h,
- the residue amount that can be dislodged from treated crops upon contact - dislodgeable foliar residue (DFR), in $\mu g/cm^2$ of foliage/kg a.i. applied/ha,
- the working time in a crop previously treated with a plant protection product (h/day), and
- the body weight (BW) of the re-entry worker in kg.

This external exposure estimate in mg a.i./kg/day is generically converted to an internal exposure estimate used in risk assessment by applying a dermal penetration factor (dermal absorption value, DA) in % of applied dose, since oral exposure studies are typically used in Europe to derive reference values (compare Section 3.1). The dermal absorption value is basically derived as a property of the product, the concentrated product or an in-use dilution according to the intended uses, which includes the recommended and/or registered application rates for every product.

Worker exposure from contact with residues on foliage is estimated using the following general algorithm:

$$Exposure\ \ dose = \frac{AppRate \times MAF \times TC \times DFR \times h \times DA}{BW \times 1000}$$

NB: the 1000 corrects the differently scaled standard units.

EFSA (2014) introduced several changes in the generic default values used in this equation. For example, the default DFR of 1 $\mu g/cm^2$/kg a.i./ha (Krebs et al., 2000), was increased to 3 $\mu g/cm^2$/kg a.i./ha (van Hemmen et al., 2002), which is a substantial increase (for details see Section 3.3.5). Moreover, the assumed body weight was decreased from 70 kg (depending on the national model used) to 60 kg, contrary to other EFSA recommendations (EFSA, 2012b), which alone corresponds to a 17% increase in exposure estimation because it is used in the denominator of the formula. These two changes increase the estimated exposure dose 3.5-times or by 250% as compared to the prior assessments. With respect to risk assessment, EFSA (2012a) measured dermal absorption

values were further generically increased from a mean to its approximate 84th percentile by addition of one standard deviation depending on experimental variation, which was recently exacerbated by further conservatisms (EFSA, 2017b) – the EFSA DA guidance changes were previously reviewed in detail (Aggarwal et al., 2014, 2015, 2019; Kluxen et al., 2019). Considering the 2012 dermal absorption modification (50th percentile as a surrogate for the mean to the 84th percentile), the estimated exposure value generically increases up to 5.9 times (3 × 1.17 × 1.68, for the fold change effect of modifying DFR, BW and DA values, respectively), equivalent to 490%, demonstrating the tremendous impact on estimated re-entry risk. This is further exacerbated when additionally, an increase of TC values is considered, as shown in Table 1, which could increase the risk almost 2000% in the grape scenario presented. While the default value changes were triggered by new data, the interaction between values, underlying assumptions, revised assessment criteria and their overall effect on risk assessment were never evaluated.

In the present manuscript, we holistically review assessment factors that contribute to potentially inappropriately overestimating re-entry exposure and thus associated risk and consider assumptions made in European guidance documents. As Table 1 demonstrates, an increase in exposure estimation might solely be driven by assumptions becoming more conservative over different guidance document generations. Subsequently, we propose risk assessment refinement strategies that - based on the reviewed information - allow a presumably more realistic estimation of the actual exposure. This prevents the need for iterative reassessment and resource used in defining further unnecessary mitigation options or data generation. Thereby, we want to enable appropriate risk management decisions based on more relevant and realistic risk estimates which allow for effective risk communication.

2. Using default values in re-entry worker risk assessment

This section discusses general aspects of using default values, which are largely ignored in the current practice, such as the multiplication of independently derived default values.

Table 1
Impact of the different European guidance document revisions on a hypothetical internal exposure assessment after an application of 1 kg a.i./ha on grapes.

Assumptions	Prior to the entry into force of EFSA (2014) and EFSA (2012)	After entry into force of EFSA (2014) and EFSA (2012)	After entry into force of EFSA (2014) and EFSA (2017)
Dermal absorption [%][a]	5.1	8.1 (mean +SD)	14 (mean +SD + 100 – recovery if < 95%)
Dislodgeable foliar residue [μg/cm^2 kg a.i./ha]	1	3	3
Body weight [kg]	70	60	60
Transfer coefficient [cm^2/h] (hand-harvesting grapes)	5000	10100	10100
Work rate [h]	8	8	8
Application Rate [kg a.i./ha	1	1	1
Estimated Exposure [mg/kg bw/day]	0.029143	0.32724	0.5656
Relation to initial estimate [%]	100	1122.9	1940.8

[a] Dermal absorption mean with a standard deviation of 3 using 7 replicates (k factor 0.92, EFSA (2017b)) in an *in vitro* assay with human skin and with a recovery of 94%. Note, for very small dermal absorption values the impact of inappropriately (Kluxen et al., 2019) adding "missing" recovery is tremendously increased.

2.1. Principles when setting default values

When using generic models to estimate exposure for non-dietary risk assessment of plant protection products, many input parameters can be set to defaults. Some default values are dependent on the physical-chemical properties of the active ingredient or product, for example those for volatility or dermal absorption. Others cover all active ingredients and products, such as the dislodgeable foliar residue value.

To derive default values, it makes sense to consider previously generated product-specific data. Until now, default values seem to be derived in complete isolation from each other (EFSA, 2014; EFSA, 2017b). However, it should be defined what amount of information warrants the derivation of a 'robust' default value. Alternatively, the amount of available information can be incorporated into the derivation as uncertainty. Neither is considered in the current EU risk assessment. It should be further noted, that the current approach only considers default point values to achieve a point exposure estimate (deterministic approach) rather than ranges or distributions of exposure (probabilistic approach).

The Precautionary Principle recommends that all available scientific data should be considered when assessing safety. If available data show that certain properties (e.g. product formulation or crop type) enable a refinement of the initial default value, that information should be used in risk assessment. Importantly, there could be several default values based on the available information, which need to be stratified based on the properties. This is currently considered for some default values, e.g. TC values, but more limited for others e.g. dermal absorption or volatilization. Generally, there is no need to set a single or a very limited number of default values when the available information allows a more relevant prediction.

When deriving default values, there are regulatory concerns associated with the approach taken and the intended meaning of a default value.

1. Typical value

A typical value (measure of central tendency) should give an exposure estimate that is commonly observed or representative for similar agents and should thence statistically lead to "common" exposure estimates. Risk derived from using a typical exposure could - for half of the estimates - theoretically decrease the amount of uncertainty allowed from the SF100-hazard characterization considerations. However, this only applies to a single exposure event. A typical exposure value considers regression towards the mean, i.e. that exposure may fluctuate over time, with higher exposures on some days and lower exposures on other days.

2. Using the possible maximum or a highest measured value

This should statistically result in highly overpredicted exposure estimates. Risk derived from using a maximum value increases the amount of uncertainty allowed from the SF100-hazard characterization, i.e. it is very conservative. That may seem health protective but might also lead to risk assessment failure of otherwise safe products with potentially worse alternatives on the market. Using maximum values may be a reasonably conservative approach when data are scarce, however, they do not incorporate any usable statistical information (also as compared to typical values).

3. A value covering most of the potential values/the distribution of an empirical data set

While "most" would be defined by risk management considerations, such values are derived either by (a) a percentile approach in combination with stratification of the data set for certain properties or (b) statistical modelling. Current exposure estimation relies on default

values that were generated with both approaches. The percentile approach assumes that a certain empirical centile relevantly and conservatively describes a distribution, which has no strong distributional assumptions. Statistical modelling always incorporates distributional assumptions and aims to incorporate precision/uncertainty, or better, compatibility with a certain range of plausible values based on data, applied statistical models and a statistical confidence for the derivation process of the interval (Gelman and Greenland, 2019) (this relates to the use of confidence intervals).

4. A prediction-based value

This is philosophically and statistically different to 3 as it considers a range of probable future values based on the historical data set and distributional assumptions and not the uncertainty associated with derived values based on past data.

Philosophical considerations and underlying assumptions, when statistical modelling is involved, tremendously affect default values derived from historical data.

It needs to be noted that in the current re-entry risk assessment approach, a single human exposure event is compared to a reference value from a toxicology study (or studies) with multiple exposure events (averaged as mg/kg bw/day). Hence, the risk assessment considers multiple subsequent daily exposures over the duration of the studies used to set reference values. For example, if a 90-day study is used to derive a reference value, this would correspond to approximately a 3-month farming season (or application/work scenario window) in risk assessment. This assumes similar exposures daily for the same task on the same crop, sprayed with the same product at the same rate, which would almost certainly never happen in practice. Short-term/acute scenarios can be considered by setting acute reference values in Europe, however, the corresponding European guidance document is not yet finalized (European Comission, 2017).

2.2. Multiplying default values

The re-entry exposure calculations used in Europe are simple multiplications of application rates and percentile default values. The issue with this approach is that the individual default values are independently derived, and no overall correlation to exposure is considered. This will, by default, result in a serious over-estimation of exposure (and correspondingly risk), as demonstrated in the following.

Fig. 1 shows the theoretical exposure calculation based on multiplication of three similar randomly-generated log-normally-distributed default values (Fig. 1A), which seems to be a common distribution for exposure values (Crowley and Holden, 2019; Korpalski et al., 2005). Multiplying just three individual default mean values substantially overpredicts the mean of a hypothetical exposure distribution and here already relates to the distribution's 80th percentile (Fig. 1B). Multiplying higher percentiles, i.e. the 95th, results in the 99.8th percentile of the resampled distribution. Multiplying more default values would additionally increase the overprediction.

This demonstrates the in-built conservatism when multiplying deterministic point estimates, which has been investigated in more detail before (Cullen, 1994). It also shows that the mean, or "typical value" of a normally distributed measure, is conservative when several means are multiplied.

3. Review of assumptions associated with worker re-entry risk assessment

In this section, we shortly review default values and some

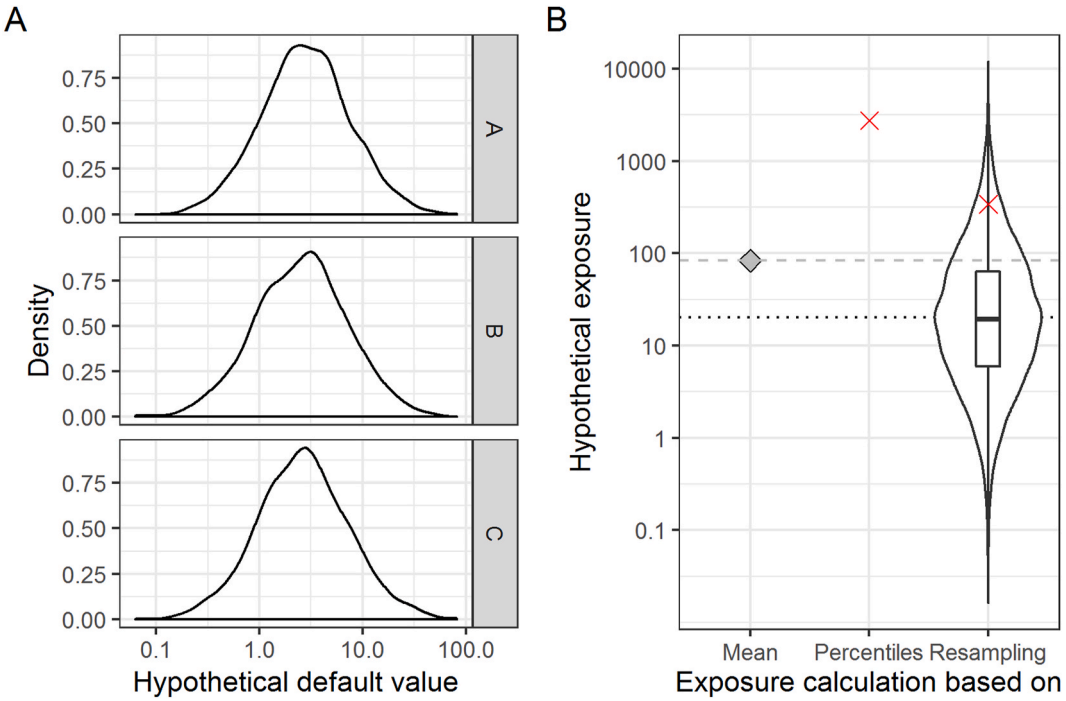

Fig. 1. Hypothetical exposure assessment using three default values from a log-normally distributed population with a mean of 1 (which corresponds to Euler's number e) and a standard deviation of 1 on log-scale (n = 1000). (A) shows the individual distributions for the hypothetical default values. (B) shows the results from a hypothetical exposure calculation (Exposure = A × B × C) based on different statistics, i.e. means, percentiles or resampling. The boxplot in B corresponds to minimum and maximum values (whisker limits) and the 25th, 50th and 75th percentile (box). For the percentile approach, 95th percentiles are calculated and multiplied. Resampling is based on taking 10000 samples with replacement from A, B and C. Here, the calculation based on generic means corresponds to the approximate 80th percentile of the resampled distribution (the generic mean is also indicated by the dashed grey line). The calculation based on 95th percentiles corresponds to the 99.8th percentile of the resampled distribution, which overpredicts the true median population exposure by about 140-fold (the location of e^3, ~20.1, is indicated by the dotted black line). Multiplying more default values based on 95th percentiles would obviously increase this ratio.

assumptions applied in the current risk assessment approach for re-entry workers in Europe.

Risk assessment is based on hazard characterization and exposure estimation. Both impact the risk assessment outcome, and assumptions in each of them are relevant when characterizing the overall coverage of uncertainties. Further, dermal absorption is an integral component of the worker risk assessment because it converts external doses into internal exposures.

While this section helps to identify areas for refinement in future guidance revisions, its major weakness is that it is only a qualitative review and assumed conservatism is not quantified. This is also the reason why conservatism needs to be applied in risk assessment *per se*, when the actual risk contribution of the single factors is uncertain. However, the current approach does not consider the propagation of conservatism, when applying conservatism on all contributing factors.

Therefore, the current manuscript focusses on a quantification of assumptions associated with external exposure, which is relatively straightforward as external exposure can be directly measured and compared with model predictions.

3.1. Hazard assessment

3.1.1. Risk driver: high dose hazard vs low dose exposure

The design of the toxicology study package considers the identification of hazard and rare events, hence, historically high doses up to the maximum tolerated dose (MTD) need to be tested. Due to the UF, the risk assessment compares by default incongruent doses of at least 100x difference. The rationale is that the UF transforms the limit value into a reference value for a potentially more sensitive human population. However, the risk assessment comparison may also become decreasingly realistic for the exposure scenario, if the human population is not more sensitive than the test species – as for example for acetylcholine esterase inhibition (Jackisch et al., 2009; Strupp et al., 2019). Also, hazard characterization of a specific effect may become confounded by using high doses, simply because physiological pathways are usually only specific in a defined concentration range – this has been termed the "Goldilocks/Lagom principle" (Leese et al., 2019). Further, if a hazard only occurs due to "metabolic imbalance", close to the maximum tolerated dose, or due to dose-driven locally high tissue concentrations, such effects are not relevant for the actual exposure scenario and the associated estimated risk.

3.1.2. Risk evaluation: hazard vs risk

Regulation (EC) 1107/2009 uses hazard-based regulation assessment criteria, which prevents the use of chemicals with certain hazard properties, based on CLP classification (European Commission, 2008) and independent of risk. This does not comply with toxicological principles and is probably politically driven due to risk aversion. While hazard-based regulation is relatively easy to perform and communicate, there might not be an overall benefit for society and it results in globally conflicting registrations for the same chemicals or products but also in different assessments in the same regions when both hazard- and risk-based regulation are applied in the same sector, such as food safety (Barlow et al., 2015).

McCarty et al. (2020) recently discussed that the concept of "inherent hazard" is flawed because also hazard depends on a certain number of molecules available to initiate an effect and toxicity is not a fixed and constant property.

3.1.3. Exposure route: animal studies vs exposure scenario

Due to animal welfare considerations, only one route of exposure, namely oral, is commonly tested in the toxicological data package for pesticide evaluation in Europe (European Commission, 2013a). In the non-dietary risk assessment, exposure is however predominantly driven by the dermal route (EFSA, 2014). Toxicokinetics differ between the exposure routes. Usually, a much higher dose is needed via dermal exposure to achieve similar AUC and C_{max} compared to oral exposure (ECETOC, 2013). And, when AUC and Cmax drive toxicity, it will accordingly be overpredicted, even when dermal penetration data are considered (because for DA only cumulative absorption is considered, but not kinetics). For acute toxicity, it has been observed that oral studies often overpredict and almost never under-predict dermal toxicity, hence the associated studies can be waived in some regulations (Creton et al., 2010; European Commission, 2013b; Latorre et al., 2019; Mielke et al., 2017; OECD, 2017; PMRA, 2017; US EPA, 2016).

3.1.4. Point of departure: limit values vs continuously increasing hazard

According to toxicological principles, the risk from most hazards is assumed to continuously increase with dose or concentration. Compensatory mechanisms, e.g. absorption, distribution, metabolism and excretion (ADME)-related, and natural biological variation in the response value (Leese et al., 2019) will however often lead to observable thresholds (see Brescia (2020) on other considerations of thresholds), not necessarily reflecting a real hazard from the exposure scenario. As the NOAEL depends on dose-spacing, which historically has been up to 10x but recommended for recent studies to be between 2x to 4x (OECD, 2008; Zarn et al., 2011), the detected NOAEL may actually under-predict an observable higher NOAEL. Conversely, the observability of a study-specific NOAEL depends on statistical power (Kluxen, 2019). However, reference values are usually derived based on a weight-of-evidence approach, which may mitigate statistical power issues to some extent. Dose-response modelling, i.e. using the BMD or its lower confidence limit (BMDL), mitigates some of the NOAEL deficiencies because it does not (usually) rely on statistical testing. However, certain assumptions on the dose-response relationship have to be applied for sparse responses and few treatment groups. Similarly, a study-specific BMDL, as an alternative to the NOAEL approach, to some extent, also depends on study design and dose-response model assumptions (Jensen et al., 2019).

3.1.5. Exposure duration: daily vs exposure according to agricultural scenario

While pesticide use may occur season-independently, e.g. indoors, most of the outdoor/field activities are restricted to the growing/harvesting season. Thus, the reference values used for risk assessment do often not correspond to the actual exposure window. For example, non-dietary reference doses/acceptable operator exposure levels (AOEL) based on 90-day feeding studies model implicitly daily oral exposure during the season, whereas the intended uses describe the actual exposure frequency. They consider the extreme but not uncommon case of only 1–2 applications in a season as compared 90 days assumed by using such a reference value. Even for big farms with sequential applications on different sites on different dates or for contract applicators an assumed constant exposure over 90-days is unrealistic. This misalignment of tested and actual exposure time is exacerbated when using chronic studies to derive AOELs. The European practice to sometimes use chronic endpoints based on identified hazards is very different to other regulatory approaches, e.g. according to US EPA (2012), "It is important that the selection of a toxic endpoint be closely matched with an expected pesticide exposure pattern to yield more accurate estimates of risk. In cases where this is not possible, assessors should acknowledge the issue and describe how this can impact the interpretation of calculated risk estimates."

Please refer to Cochran and Ross (2017) for additional considerations.

3.2. Dermal absorption

In Europe, estimated or measured external doses are compared to an internal reference value, the AOEL, by using a dermal absorption factor. Contrary to OECD GD 156 (2011), which proposes a worst case value of 100%, the four available European default values are based on

formulation type properties (EFSA, 2017b). A refinement is possible by conducting product-specific dermal absorption studies.

3.2.1. Penetration vs absorption

Human dermal absorption can be estimated using animal models, similar to *in vivo* studies on ADME by other routes of exposure (WHO 2006). The test item is applied to animal's, usually rat, fur-free-clipped intact skin under a protective cover. Excreta and exhaled air are collected for several days after initial exposure. The amount of test item is analysed in all samples and carcass. Dermal absorption can also be estimated with *in vitro* penetration models using human or rat skin (other skin types are also used, e.g. pig skin). In a diffusion cell, a compound penetrates from an upper donor chamber through outer skin layers into a lower fluid filled receptor compartment, which is used as a surrogate for the systemic body compartment, in either a static or a flow-through system. Samples taken from the receptor fluid over time allow kinetic investigations. Migration through the skin is usually followed for a maximum of 24 h after the start of exposure, because the integrity of the skin sample may be susceptible to deterioration over longer periods. However, today, models are also routinely run longer than 24 h, compare e.g. Gunther et al. (2019). Numerically and due to the availability of human skin, *in vitro* assays are today's standard tool for addressing dermal absorption of pesticides. Further, EFSA (2017b) discourages the use of animal studies, due to animal welfare concerns and the *in vivo* studies' overprediction for human risk assessment, because of differing skin morphology. Currently, for Europe, *in vivo* studies are only used in conjunction with human and rat *in vitro* studies to triangulate potential human absorption, i.e. the so-called triple-pack approach, to account for *in vitro-in vivo* differences.

Thus, dermal absorption estimates, as commonly used today, are actually surrogates from penetration assays that disregard *in vivo* ADME effects. This is illustrated by the de-facto standard approach of using radioactively-labelled compounds, which allows a straightforward scintillation analysis and increases sensitivity but ignores metabolic processes.

3.2.2. Relative vs absolute amount

Dermal absorption is in principle a passive diffusion process where the compound of interest passes the skin barrier by either one or several of the possible passage routes, i.e. transcellular, intercellular or trans-appendageal (i.e. bypassing the skin via hair-follicles, skin glands or sweat ducts), depending on the compound's physical-chemical properties. The current model approach does not consider actual exposure, in terms of skin loading $\mu g/cm^2$, but relative absorption expressed as the fraction (%) penetrated of the applied dose. This percentage approach has several implications in terms of being conservative. For actual exposure (absolute amount absorbed), the dose response is usually as expected, i.e. the higher the applied dose the higher the actual amount absorbed, unless the diffusion capacity is rate limiting and thus the amount applied is in excess to what can diffuse over time. For the relative fraction approach, the relationship is however usually inverse, i.e. the lower the applied dose per area the higher the penetrating fraction. In theory this could go up to 100% but the question remains, 100% of what? It is an obvious difference in terms of exposure if one considers 100% of a single molecule as compared to 100% of 1000 or 1 million molecules per skin area. If one now considers that this is a common risk assessment approach to use the dermal estimate for the lowest concentration as a surrogate for higher concentration, this may result in a significant overestimation of actual exposure.

For use of dermal absorption estimates from spray dilutions to assess dermal absorption of dried residues there is another aspect of penetrated fraction that may play a crucial role in terms of overestimation. To enable diffusion, the dried compound transferred to the skin surface must solubilize at the barrier of the skin matrix. It is plausible that this solubilization will likely affect only a fraction of the transferred dose to the worker's skin. This solubilization step is already considered when the compound is applied to skin in a solubilized form, as is the case for spray dilution testing.

3.2.3. Product property vs exposure scenario

Dermal absorption studies do not correspond to any actual skin exposure dose – at least not by default or intention and based on the currently requested design for regulatory purposes. The studies correspond rather to concentrate and in-use concentration (or slurry) properties. The choice of test concentrations is not driven by exposure-based dermal loadings and distribution on human body. While there are approaches being developed on how such doses could be set with regard to the worker scenario and dried residues (Morgan et al., 2020), many elements of the standard DA studies themselves do not correspond to the scenario, as reviewed in this section.

3.2.4. Dried residues

According to good agricultural practice, workers may only enter a field after the application solution has dried (or a defined re-entry period based on the risk assessment). Also, the current worker exposure assessment according to EFSA (2014) only allows estimations of the application once the solution has dried. While dew or rain may have re-wetted the dry residues, they would presumably also affect the amount of residues left for worker exposure due to run-off.

The available data suggest that using DA values for in-use dilutions generically overpredict dried residues dermal absorption (Aggarwal et al., 2019; Clarke et al., 2018), e.g. 2–3 times based on the percentage of the absorbed dose, when the same amount is transferred as dried residues as tested in the corresponding in-use dilutions. In practice, compared to skin contact with leaves covered with wet spray concentrations less compound will be transferred to the skin when the leaf residues have dried. However, the EFSA DA guidance (2017b) proposes "that the appropriate dermal absorption value for exposures to dried dispersed residue should be the higher of the values for the concentrate and the in-use dilution" which in reality is almost always the in-use dilution.

For water-soluble products, the concentrate is most closely related to dried residues as it lacks the water used for preparing the application solution. Dried residues are by definition dry and cannot penetrate into the systemic compartment. However, we usually assume that sweat or dew provides a certain amount of water, which is considered when conducting DA assays. For example, for solid products, concentrates are tested as 1:1–1:3 slurries with water. The DA from solid products is shown to be very low (Aggarwal et al., 2014, 2015). However, relative DA is driven by decreasing absolute amounts, hence one could argue, ignoring the physical-chemical properties, that this makes dry residues similar to highly diluted in-use dilutions. Taken together, neither approach seems appropriate.

When only worker exposure is the risk driver, it would hypothetically make sense within the risk assessment framework, to only test a concentrate to prevent having to produce and use the in-use dilution DA value for worker exposure, since concentrates usually have lower relative DA values. This illustrates the highly unrealistic nature of the approach.

As a pragmatic mitigation and to acknowledge the conservatism of using concentrate or in-use dilution values for worker exposure one could use a fraction of the in-use dilution DA value ($\times$ 1/3) or a multiple of the concentrate value ($\times$ 6), whatever is less (multipliers based on rounded means of ratios from dry residue and concentrate or in-use dilution relative DA data) based on the data presented in Aggarwal et al. (2019). Such an approach is however not yet introduced in the EFSA guidance on dermal absorption.

3.2.5. Study length vs exposure time

In vitro dermal absorption studies are usually conducted with 6-10-h exposure after which the skin sample is washed, the sampling of the receptor fluid is continued until 24 h after initial exposure.

Early decontamination of exposure skin samples has a dramatic effect on cumulative dermal absorption (Clarke et al., 2018). Thus, for short duration inspection activities, e.g. in cereals or other field crops, the exposure duration in these studies highly overestimate the actual worker exposure duration.

3.2.6. Cumulative absorption vs systemic dose

Currently, the internal exposure estimate is usually not refined by considering other toxicokinetic information available in the regulatory study package, for example systemic half-life or kinetic information from the dermal absorption studies.

Standard assays conducted according to OECD TG 428 consider a 14-18-h follow-up after a 6-10-h exposure period. Hence, the dermal absorption estimate from such studies assumes cumulative 24-h exposure, i.e. no consideration of ADME effects during this time. Just using absorption kinetics thus overestimates internal exposure by default, since the exposure model assumes that total daily external exposure starts at t = 0 of the exposure scenario. The cumulative dermal absorption value assumes that everything that actually penetrates only after 24 h would penetrate at t = 0, i.e. at initial exposure. Hence, the typical lag-phase and chronologically increasing absorption are ignored. Also, metabolic detoxification and clearance by excretion are not taken into account. All these factors decrease the overall area under the internal exposure over time curve, see Fig. 2. The figure depicts the conceptional exposure model and some factors that affect the internal dose. In reality, only a fraction of the internal exposure dose is available at any given time point, as internal exposure is dynamic. The latter obviously also applies to reference values, hence, average internal exposure or PBTK modelling may be better suited for more relevant risk assessments.

3.2.7. Pro-rata extrapolation

According to EFSA (2017b), if a concentration is used in the risk assessment that is more diluted than the experimentally tested concentrations for dermal absorption, it is suggested to consider whether the DA values have to be adjusted pro-rata for the use scenario. The data and basis for this approach is weak, compare Figure B.3 in EFSA (2017b) or other sources (Aggarwal et al., 2014, 2015), and it is acknowledged to be conservative by EFSA. However, as it is recommended in the EFSA guidance it is generically applied in Europe (experience of the authors that are involved in European registration processes).

The pro-rata approach has tremendous implications because dermal absorption values are used as a fraction to modify the external dose, hence a generic modification directly affects the estimation of the internal dose. One can construct a hypothetical scenario, where a small absolute amount is applied, as compared to other uses, in a very low-concentrated in-use dilution, which makes the pro-rata modification necessary. As the pro-rata approach increases the DA value, this scenario may suddenly become the risk driver as compared to higher concentrated use rates. However, this only occurs when a pro-rata adapted dermal absorption value is derived for a low application rate but combined with the risk assessment for a high application rate, assuming that water volume does not change. For all other cases, the linear pro-rata increase in absorption balances out to the linear application rate decrease for dermal route estimates in re-entry worker assessment (same applies for dermal exposure scenarios for bystander and resident). Low application rates with concurrent high pro-rata adapted dermal absorption estimates can only become a risk driver scenario for the operator, where the application rate dependency is log based as compared to the linear based dermal absorption adoption.

Recently, the pro-rata modification has also been used by authorities to increase the DA values used in worker exposure, i.e. when the workers are exposed to dry residues of the sprayed dilutions also when workers enter the treated areas days after application. This approach seems to be conflicting with the EFSA risk assessment guidance itself, see above on the requirement. The rationale has been put forward that the DA value, for example, for operator exposure needs to be increased pro-rata; then the extrapolated DA value becomes the highest available DA value. The key driver of worker exposure to dry residues is the absolute amount applied to the re-entered area. It is obvious that different dilutions of the same absolute amount applied on the same area result in the same absolute amount of dried residues, however, only if other factors such as run-off and distribution within the crop space are ignored – which are in-fact not considered in the current model assumptions and cannot be methodically considered within the assay (compare discussion regarding the DFR value). Hence, the pro-rata approach on in-use dilutions is very conservative for dried residues. The available data

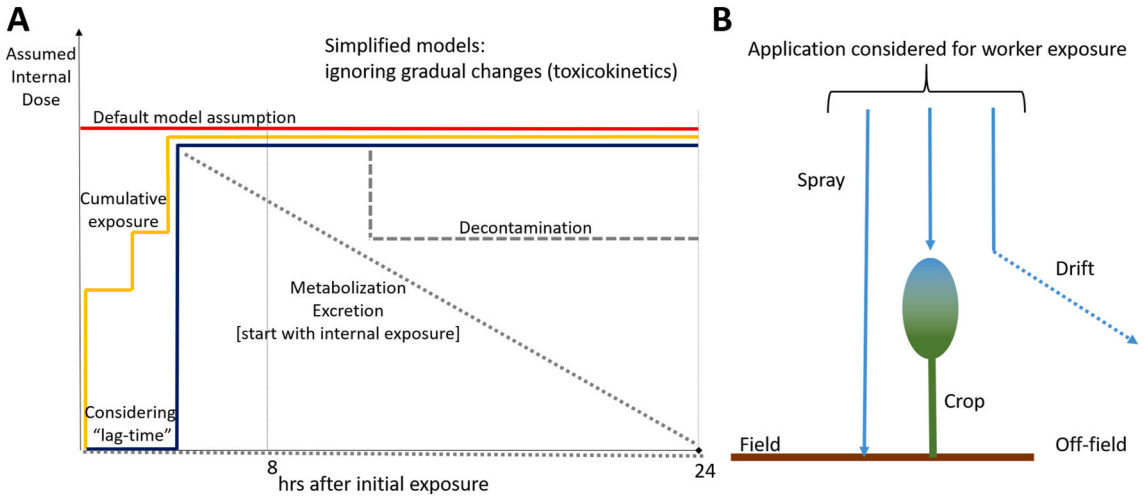

Fig. 2. Conceptual European worker exposure model assumptions. A) The default model maximizes internal exposure. Plotted is the internal dose after initial exposure against the time after multiple exposure events within 8 h over 24 h in total. The default model assumes full work-day exposure at the time of first exposure. Cumulative exposure will, however, only happen over time, either by multiple exposure events or due to a lag-time of diffusion/absorption through the skin (here symbolized by a 4-h lag and a single exposure event at 4 h). Internal dose is further affected by ADME/toxicokinetics and decontamination procedures, which is indicated by the dotted and dashed grey lines in the plot, but actually results in gradual/dynamic changes of the internal dose, which is ignored here. It is obvious that the models result in vastly different internal exposures. Usually only the default model is used in pesticide risk assessment in Europe. Dermal absorption studies often report a certain lag-time until exposure of the systemic compartment surrogate, such a lag-time also leads to reduced internal exposure estimates. B) Assumptions disregard conceptual model deposition patterns, e.g. as used for risk assessment of bystanders or other sections such as efficacy, ecotoxicology and environmental fate.

(Aggarwal et al., 2019; Clarke et al., 2018) also suggests that using DA values for in-use dilutions even without pro-rata modification generically overpredict dried residues dermal absorption, as described above.

3.2.8. Considering >24-h dermal absorption vs daily reference values

One of the major changes in the original EFSA DA guidance (EFSA, 2012a) and its revision (EFSA, 2017b) was to consider 100% of the residue in the *stratum cornea* to contribute to systemic exposures, based on kinetic parameters, even if the residue did not penetrate into the systemic compartment surrogate "receptor fluid" within 24 h. The reasoning is that it could potentially penetrate later which ignores the fact that risk assessment is conducted with 24-h reference values. The argument is curious because it is used in conjunction with cumulative absorption, where such temporal/kinetic considerations are ignored.

The amount remaining in the *stratum corneum* either gets desquamated over time or enters into the systemic blood flow very slowly, leading to no considerable increase in AUC or C_{max} and therefore presumably not driving systemic toxicological effects. This is further underpinned by ignoring clearance aspects.

Therefore, compounds that do not penetrate efficiently but remain in the *stratum corneum* are disproportionally penalized with overpredicted 24-h relative dermal absorption values.

While human data shows that systemic exposure correlates best with the *in vitro* receptor fluid values (Lehman et al., 2011), this did not affect the approach for deriving potentially more relevant dermal absorption values. Hence, it would make sense to discuss within the regulatory community how this assumption can be countered by data, even if the *a priori* assumption is implausible with respect to the risk assessment framework. Two immediate approaches appear suitable: the ratio of residue in *stratum corneum* to receptor fluid gives an indication of the relevance that the *stratum corneum* residue contributes to the dermal absorption estimate. If the residue is a multiple of the receptor fluid value, one may scrutinize its relevance. Another approach would be to produce data showing that the amount in the *stratum corneum* does not relevantly contribute to the post-24-h receptor fluid recovery; as such longer-term studies appear to be robustly feasible today (Gunther et al., 2019).

3.2.9. "Missing material" = absorbed material

According to EFSA (2017b), relative dermal absorption estimates below 5% should be adjusted for "missing material" if their mean recovery is below 95% in the studies. The reasoning is that on the one hand "modern analytical and pipetting techniques" would enable high recovery in studies and on the other, an uncertainty argument, "if 9% of the test material is unaccounted for, there is a high degree of uncertainty surrounding a proposed dermal absorption value of 1%". Neither claim seems to be data-driven, according to the EFSA dermal absorption database, or recent developments in the analytical field, and both have been recently refuted (Kluxen et al., 2019, 2020). Relative dermal absorption and recovery do not correlate in the range that is acceptable according to OECD test guideline 428 (OECD, 2004). "Missing material" is actually "insufficient dosing". The radiolabel recovered can only reflect the amount of radiolabel initially applied. Hence, the approach does not account for uncertainty but according to EFSA guidance (2017b) leads to grossly overestimated dermal absorption estimates for exactly those compounds that do not efficiently penetrate into the systemic compartment.

3.2.10. Accounting for variability vs variable exposure

In contrast to the evaluation of ADME studies, EFSA (2012a) introduced a secondary evaluation criteria to assess specifically dermal absorption studies. If the variability in the results, measured by standard deviation, for relative absorption, which is compounded by adding various fractions of residues recovered in the assays (*stratum corneum*, remaining skin, receptor fluid), exceeds a predefined threshold of 25% of the mean absorption, the standard deviation is added to the mean to derive the dermal absorption estimate. However, this procedure disconnects the estimate from the applied dose and other parameters measured in the study, e.g. washed-off fraction. As both mean and standard deviation are biased by extreme values in either direction but is only accounted for in one direction in this process, one may end up with grossly overpredicted absorption estimates. Especially for low absorption compounds, single high values may increase the standard deviation to be higher than the mean value.

Consider the following examples: The data set of [0.02, 0.01, 0.01, 1] has a mean 0.26 and a standard deviation of 0.49. Hence, the estimate ends up being [0.26 + 0.49 =] 0.75. None of the measures seems to represent the data set well, but it is very problematic to exclude a potential outlier in regulatory practice, because the amount of evidence necessary for this is not well-communicated or harmonized.

The random values [0.1, 0.8, 1] have a mean of 0.633 and a standard deviation of 0.473. The modified dermal absorption estimate will be 1.1, which is higher than all individual values.

EFSA (2017b) modified the approach to account for the available information used to estimate dermal absorption, where an increased number of observations lead to a smaller fraction of standard deviation added to the mean. This will decrease the upper confidence limit associated with the mean. In our previous extreme examples, the estimates would increase due to the limited observation numbers.

While the reasoning in the DA guidance revision is given as a refined estimation based on the amount of data, the 2012 guidance argues that studies with increased variation may be unreliable; "If there is significant variation between replicates consideration should be given to using a value other than the mean or rejecting the study entirely", which conflicts with OECD TG 428.

It needs to be noted that the approach does not "take account of variation" but specifically increases dermal absorption estimates due to variation. In reality, variation, especially when skin donor-driven, means that the DA value may be underpredicted for some but overpredicted for other individuals, but only if other factors driving dermal absorption are ignored.

Overall, it may be discussed whether mean and standard deviation represent the dermal absorption data well at all. According to EFSA (2017b) they do not (see Appendix B in the guidance document) - but are still recommended to be used. Especially for compounds with predominantly low absorption, as in the first example, other measures, such as the geometric mean, may be more appropriate (compare mean of 0.26 to geometric mean of 0.04 for [0.02, 0.01, 0.01, 1]).

3.2.11. Studies conducted with dorsal or breast skin vs hand exposure

In vitro dermal absorption studies for worker risk assessment are usually conducted with human donor skin, obtained from body donors or cosmetic surgeries. Due to the higher supply from cosmetic surgeries, the available skin is usually abdominal, dorsal or breast skin.

There are morphological differences between abdominal, dorsal and breast skin when compared with skin mostly in contact with crops: hands. Absorption may also be affected by lipid content in the *stratum corneum* (Elias et al., 1981). While this is universally acknowledged and an effect on dermal absorption is plausible, this was not systematically investigated and available data are scarce (Lev-Tov and Maibach, 2012). Investigated absorption differences mostly rely on a few studies (Feldmann and Maibach, 1967; Maibach et al., 1971; Rougier et al., 1986). A recent review on impact of anatomical location on dermal penetration in man (Bormann and Maibach, 2020) also considers newer studies on transdermal drug delivery identifying forehead, neck and genitals as the most vulnerable skin regions. However, as these drug delivery studies focus on penetrable anatomical skin locations applicable for drug delivery, little or no new information is available for the comparison of the tested anatomical site abdomen or back in relation to those sites most relevant for pesticidal exposure during re-entry activities, i.e. hands, forearm and lower legs. It has been previously proposed to use such differences to derive correction factors for exposure estimations (Guy

and Maibach, 1984, 2002), which is not applied in practice. However, the dermal absorption assay conduction has progressed continuously and is now harmonized and easily accessible by available *in vitro* test guidelines, which would make a systematic investigation worthwhile.

3.3. Exposure assessment

The calculation for worker exposure considers the full application rate to contribute to the dislodgeable foliar residue amount, disregarding the crop type, leaf morphology, leaf development stage, etc. - see calculation in the introduction. However, when one considers the conceptual exposure scenario (Fig. 2B), it is obvious that only a fraction of the applied rate can actually become available for worker exposure. This is reflected by risk assessment assumptions for other target populations, such as bystanders, or in other sections such as efficacy, ecotoxicology and environmental fate.

For example, a certain amount of a.i. is potentially lost by direct spray drift and volatilization. As reviewed below, this may amount up to 50% of the applied dose, which implies that this amount is not available for worker exposure. Also, the dislodgeable amount for systemic pesticides is reduced by definition when a certain amount has been absorbed in the plant, which is desired for many pesticides. Ecotoxicology risk assessment is mainly driven by deposits on the ground and considers plant interception parameters in the calculations. Environmental exposure and degradation are also not considered by default. Conversely, when re-entry intervals are set, the calculations based on DFR/degradation studies, rainfall during study conduct period often affects the acceptability of the results.

All of this is ignored in worker risk assessment, which further compounds the overestimation within this scenario. It would be reasonable to assume that a holistic approach to risk assessment, which consolidates assumptions between sections, directly affects the amount of overestimation in the individual assessments. Probabilistic approaches could also incorporate information on environmental conditions, e.g. rainfall, which would dramatically reduce average worker exposures and is a common exposure scenario.

3.3.1. Risk envelope vs agricultural scenario

The selection of the critical GAP (Good Agricultural Practice) use should be based on actual uses in the field. However, the presentation of a GAP range includes different application rates and product dilutions (relevant for selection of critical dermal absorption estimates). Usually, the spray dilution per application rate is driven by crop habitus and application equipment, thus, each application rate has usually a dedicated dilution range. Meanwhile authorities often request to combine the critical, i.e. maximum, application rate with dermal absorption estimates for the highest dilution in the GAP range, i.e. which is related to the lowest application concentration only, assuming water volume stays the same when changing the application rates. If not properly covered by testing, pro-rata adoption for dermal absorption is applied. By this approach a further safety factor is added in each “risk envelope” scenario, due to the difference in concentration between highest and lowest application rates. This risk envelope is, therefore, not reflecting the critical GAP but a combination of 2 GAPs within a GAP range, by this, resulting in addition of an artificial calculation safety factor. This approach seems to be conflicting with the EFSA risk assessment guidance itself and also with the EU risk envelope guidance (European Commission SANCO, 2011) that requires to identify and assess the critical (actual) GAP.

3.3.2. Multiplication of independently derived high percentile default values or division by low-percentile defaults

Default values used in non-dietary risk assessment are currently derived independently of each other. Since the multiplication of point default values overpredict the resulting risk distribution, this is a very conservative approach; refer to Section 2.2 and Fig. 1.

Since exposures vary over a working day, for example because foliar residues vary over a treated field, which is the reason why multiple samples are tested to determine residues, high and low exposures.

3.3.3. Full exposure dose at beginning of exposure

While re-entry exposure continuously increases over time and may reach an equilibrium in re-entry scenarios (Ross et al., 2000), the model assumption conversely assumes an initial full exposure at $t = 0$ which is then maintained throughout the exposure period. By extension, it assumes exposure continues for the rest of the 24-h day in the risk assessment based on the daily reference value and daily thereafter corresponding to the reference value. If one only considers that exposure cumulatively increases over the exposure period to a maximum and no other toxicokinetic factors apply, it is obvious that the internal dose is systematically overpredicted, as shown in Fig. 2A. In reality, a full initial level of exposure is not maintained throughout the day and internal dose changes dynamically over time due to ADME processes, which is investigated in the requisite ADME studies for the a.i.. Hence, it is in principle possible to refine internal exposure estimates by kinetic modelling – either based on the model assumption or defined cumulative exposure patterns for the various scenarios, e.g. continuously increasing exposure vs time-defined bolus events. Such an approach has been explored for example for e.g. haloxyfop (Cooper et al., 2018).

3.3.4. Body weight

EFSA model assumptions consider 60 kg body weight, in contrast, according to EFSA Scientific Committee (2012), 70 kg should be used as a default value because it is a closer approximation to the mean body weight of the EU adult population. Since the exposure is divided by body weight (see Introduction), a lower body weight default value increases the exposure estimation per kg.

US EPA assumes a mean body weight of 80 kg (combined adults) and 69 kg (female body weight to represent average body weight of women of child-bearing age, assumed to be ages 13 through 49), the latter being similar to the EFSA Scientific Committee proposal. Data provided in the Exposure Factor Handbook for ages 11 through <50 were averaged to represent this life stage (US EPA, 2011).

3.3.5. Dislodgeable foliar residues

The DFR is the key driver of worker exposure estimation, because it links the amount of dislodgeable residues from crops with the application rate. As reviewed above, the amount applied and thus available for dislocation is overestimated by generic model assumptions. Further, the DFR only considers the absolute a.i. amount applied, independent of the used dilution rate. The model is thus not affected by applied water volume, while water volume affects the a.i. distribution on the crop. Thus, there is also a contradiction to recent authority approaches of using pro-rata correction for dermal absorption values for dried residue worker exposure (compare Section 3.2). The reasoning for the pro-rata approach suggests that less absolute amount would be considered for exposure (because the absolute amount within the model is not affected by dilution), which does however not correspond to the DFR value describing the exposure.

The DFR value is not stratified by crop type or crop type group due to insufficient data (EUROPOEM II, Annex 9). However, the DFR could be assumed to be directly affected by considered crop type and growth stage if one compares, for example interception assumptions in ecotoxicology and environmental fate predictions (EFSA, 2017a; van Beinum and Beulke, 2010) as discussed below.

When DFR studies –or efficacy and residues trials– are conducted, it is crucial to sample at multiple field locations due to variation in spray homogeneity. Hence, the deposited dose within the field varies, from which follows that also the potential dose that can be dislodged by workers varies over the exposure period. This illustrates the conservatism of using high percentiles to derive DFR default values. Therefore, deriving mean values would seem to be appropriate.

The current default DFR in the EFSA guidance risk equation comes from the EUROPOEM II project (van Hemmen et al., 2002), which compiled the at that time publicly available scientific literature and authority generated data in order to extract data on initial DFRs, without using proprietary studies. The database comprises 55 studies from 1958 to 1999, including 46 active substances and 28 crop types. EUROPOEM II suggested that for a highly conservative assessment of the initial DFR (DFR0), in a first tier assessment, 3 μg/cm^2 active substance on foliage, which is about the 90th percentile of the distribution, can be taken as a default value when no relevant/appropriate data on leaf area index can be used to estimate worker dermal exposure. According to the publication, the full study list was "not presented due to size", which makes derivation of DFR value intransparent.

CropLife Europe (CLE, former European Crop Protection Association, ECPA) is currently analysing DFR data from member companies, i.e. >390 study reports and >1250 datasets, in order to explore bridging options and drivers of DFR (Blaschke, 2020). This project could obviously result in a revision of the current default DFR in the future.

3.3.6. Foliar half-life

The default value for foliar half-life (DT_{50}) is used to consider post-application intervals for worker exposure estimations, i.e. re-entry after a certain time period during which the foliar residue degrades and reaches an acceptable level. The DT50 value of 30 days for worker exposure relies on 277 substances from the USDA ARS pesticide properties database and is a 95th percentile value. This dataset contains a high proportion of values of 30 days or higher, about 13% (i.e. for 36 a.i.s), which appears as a clear spike in a kernel density graph of the data (graph not shown). The fraction below 30 days is well characterized with a geometric mean of 4.7 days. Out of 36 a.i.s, only 8 (fluometuron, formetanate hydrochloride, mepiquat chloride, metalaxyl, metsulfuron-methyl, pendimethalin, prochloraz, thiabendazole) are currently registered for use in EU; all of these have half-life of 30 days according to these data and constitute < 3% of the whole database. It has to be noted, that although EFSA refers to this database for setting the default DT50, it is not accepted to use the same database in order to choose a substance-specific DT50 value.

The European risk assessment for birds and mammals considers 10 days as a default DT50 (EFSA, 2008). This value relies on a dataset from Willis and McDowell (1987), comprising 450 DT50 values of which 81 are agricultural pesticides. According to EFSA (2008), due to the time schedule of sampling in the original studies, the authors expect that many of the half-lives may be overestimates, additionally many very stable substances such as organochlorines are contained in the database (thus distorting the estimation). The default values for FOCUS surface water modelling in environmental fate was accordingly also changed to 10 days to specifically harmonize assumptions between risk assessments.

This illustrates that non-dietary risk assessment assumptions are derived in isolation from other sections, even though many of these processes are inter-connected.

3.3.7. Transfer coefficient

The transfer coefficient (TC) quantifies the extent of foliar contact during re-entry depending on a specific work scenario. EFSA (2014) relies on percentile values from a limited and unbalanced number of studies reported in the EUROPOEM II project (van Hemmen et al., 2002).The percentiles vary between crops: 75th percentiles for hands and body were used for vegetables, 90th percentiles for hands and body were used for tree fruits, and the 75th percentile for hands and 90th percentile for body were used for ornamentals.

Additionally, the weight of specific percentile values depend on the distribution of exposure on body. An example of exposure distribution and corresponding percentiles: e.g. 75th-ile value for hands contributes to 88% of exposure; 90th-ile value for body contributes to 12% of exposure.

3.3.8. Crop interception vs 100% applied dose

Only a fraction of the applied amount ends up on the intended target, which is a key issue for efficacy. Accordingly, only a fraction of the applied amount is available for re-entry worker exposure. While this is ignored in the current exposure estimation, it is considered in other assessments, e.g. environmental fate and ecotoxicology.

The interception for all crops is considered to be between 0% (pre-emergence) up to a maximum of 90%, even with full foliage (BBCH99), in the generic guidance for tier 1 FOCUS groundwater assessments (FOCUS, 2014). For example, for apples, the interception rate is only 65% even when considering a full canopy.

3.3.9. Plant absorption

When pesticides have systemic efficacy, they are by definition not fully available for worker exposure. Also pesticides that are not intended to have systemic efficacy may still penetrate into the crop and reduce the available applied dose (Yang et al., 2017) for worker exposure.

3.3.10. Volatilization

Any fraction lost by volatilization cannot remain as dried residue. While inhalation exposure from volatilization and/or drift loss is not considered in the worker risk assessment, the recent exposure studies shown in Section 4, which include also measurement of inhalation exposure, indicate that inhalation exposure is negligible. This is not surprising as a field is spread over a large area with changing wind direction and speed, and it is not a confined space (Fig. 2B).

4. Model realism based on available data

EFSA (2014) uses a regression-based approach on measured exposure data for operator exposure. A similar dataset was not available for worker exposure because they have never been a population under increased calculated risk. A further reason is the logistic and economic impact of worker exposure studies. There are however exemplary studies available for certain crops and products to overcome the above discussed exaggerated conservatism in the risk assessment model applied. For instance, due to the revised TC value in grapes (EFSA, 2014) and the lack of appropriate publicly available data, a collection of proprietary industry studies in grapes have been made available to European authorities, which based on their review allowed a proposal for revised lower TC values in grapes (HSE et al., 2020),[1] i.e. the "BROV" (Bystander Resident Orchard Vineyard) dataset. While the implementation of the BROV data in the EFSA guidance revision is uncertain at the time the manuscript is being prepared, it shows that robust data can be successfully generated by task force projects in Europe, which is similar to projects in the United States. There is, however, currently no legal framework for EFSA to actively participate in such projects. Other studies for relevant crop groups with re-entry tasks – scouting and inspection in field crops (cereals) or maintenance activities in orchards (apple) – have also been reviewed for this manuscript.

In the following, the measured exposure estimates from 16 available GLP-compliant worker studies (Table 2) are compared with generic model default values assumptions (Table 3) and calculations. The studies were collected from CLE member companies, which have used the studies in regulatory processes. The studies are only briefly compared in Table 2, which gives the key parameters used in the corresponding generic model calculations.

In exposure measurement study reports, for every location and crop, nominal application rates have been stated, and as such they have been used in the calculations according to the EFSA model (the same way it would be done in the registration process). Consequently, for specific locations, application rate regimes and crops, the calculated exposure

[1] Available at croplifeeurope.eu/media/news/bystander-resident-orchard-vineyard-brov-re-entry-project-report/

Table 2
Key characteristics of the worker re-entry exposure study dataset.

Study	Reference	Application rates and details
		Grapes
a	000082110	10 × 1.44 kg a.i./ha; or 10 × 1.52 kg a.i./ha Interval: 7 days REI: 1 day (when spray has dried) Number of workers: 12 Activity: pruning, tying back
c	000082535	1 × 0.75 kg a.i./ha REI: 1 day (when spray has dried) Number of workers: 12 Activity: harvesting
n	2016/1222400[a]	1 × 1 kg a.i./ha; or 1 × 0.8 kg a.i./ha REI: 2 h (when spray has dried) Number of workers: 12 Concurrent DFR measurement per study site, data part of the BROV WoEG Activity: pruning
o	2019/1043774	1 × 1 kg a.i./ha; or 1 × 0.8 kg a.i./ha REI: 2 h (when spray has dried) Number of workers: 16 Concurrent DFR measurements per study site Activity: pruning, tying up
		Apple or pome fruit
b	000082109	9 × 1.44 kg a.i./ha + 1 × 1.80 kg a.i./ha Interval: 7 days REI: 1 day (when spray has dried) Number of workers: 12 Activity: thinning, pruning
k	2015/1113647	3 × 0.45 kg a.i./ha Interval: 7–10 days REI: 1 day (when spray has dried) Number of workers: 12 Concurrent DFR measurements per study site Activity: thinning fruits, pruning trees
l	2011/1122477	3 x 0.525 + 3 × 0.35 kg a.i./ha Interval: 4–7 days REI: 1 day (when spray has dried) Number of workers: 12 Activity: thinning fruits, pruning trees
m	2020/200620	6 × 0.35 kg a.i./ha Interval: 4–6 days REI: 1 day (when spray has dried) Number of workers: 12 Concurrent DFR measurements per study site Activity: thinning fruits, pruning trees
d	ACI19-011	2 × 0.5 kg a.i./ha Interval: 28 days REI: 2 days Number of workers: 12 Activity: thinning/pruning
		Stone fruit
p	S 15-03770	3 × 0.05 kg a.i./ha; or 4 × 0.05 kg a.i./ha Interval: 10 days REI: 3 days Number of workers: 12 Activity: harvesting
		Maize
e	VV-470176	1 × 0.495 kg a.i./ha REI: 1.5–2.75 h (when spray has dried) Number of workers: 12 Concurrent DFR measurements per study site Activity: crop inspection/scouting
		Wheat
f	VV-415297	1 × 4.0 kg a.i./ha REI: 1–2 h (when spray has dried) Number of workers: 12 Concurrent DFR measurements per study site Activity: crop inspection/scouting
g	2015/1076589	2 × 0.125 kg a.i./ha (a.i. 1) Interval: 10–26 days REI: 2 h (when spray has dried) and 1 day Number of workers: 12 (for each timepoint)

(*continued on next page*)

Table 2 (*continued*)

Study	Reference	Application rates and details
h	2015/1076589	Concurrent DFR measurements per study site Activity: crop inspection/scouting 2 × 0.17 kg a.i./ha (a.i. 2) Interval: 10–27 days REI: 2 h (when spray has dried) and 1 day Number of workers: 12 (for each timepoint)
i	2014/1037490	Concurrent DFR measurements per study site Activity: crop inspection/scouting 2 × 0.125 kg a.i./ha (a.i. 1) Interval: 20–22 days REI: 2 h (when spray has dried) Number of workers: 12
j	2014/1037490	Concurrent DFR measurements per study site Activity: crop inspection/scouting 2 × 0.125 kg a.i./ha (a.i. 2) Interval: 20–22 days REI: 2 h (when spray has dried) Number of workers: 12 Concurrent DFR measurements per study site Activity: crop inspection/scouting

[a] included in the BROV WoEG report.

estimates in μg/kg bw/day have been derived. In the studies, total external cumulative dermal exposures (actual exposure body; potential exposure hands (i.e. assuming normal workwear and bare hands) for every individual worker have been reported, i.e. exposures in μg/person have been adjusted for worker body weight to derive normalized μg/kg bw/day values. The measured individual external dermal exposure values were compared with exposure estimations achieved according to EFSA guidance, and ratios between calculated exposure estimates and measured ones have been derived.

Fig. 3 shows the individual model to study exposure ratios by study and crop. The figure shows that the actual exposure is never underpredicted by the generic model. The ratios conversely demonstrate the scale of overprediction when using the EFSA model: the current EFSA model (2014) overpredicts potential exposure about 50 times on average (median of study medians).

Acc. to EFSA (2014), "The main routes of exposure during post-application activities are dermal and inhalation […]" and "[..] *in many cases, inhalation exposure is likely to contribute less to total potential exposure than that by the dermal route* […]". In some of the available worker exposure studies, also data on the inhalation exposure are available. In these studies, the mean inhalation exposure is indeed only a fraction of the dermal exposure, i.e. < 1% on average from potential and actual dermal exposure in the available studies. This confirms an insignificant contribution to the overall exposure; hence it will not be further discussed in the current manuscript.

The reason for the model overprediction can be partly explained here by reviewing the body weight data (Fig. 4) of the workers that participated in the studies. While some workers have lower body weights than the default value of 60 kg, which is expected for such distributions, most of the workers have higher body weights. Hence, the mean prediction is biased for most workers. It should be noted that the density graph Fig. 4c, shows two peaks which potentially indicates that both male and female workers are present in the study population. Unfortunately, this was not documented in all studies, hence this could not be further investigated.

Notably, the other factors discussed in Section 3 affecting external exposure, *i.e.* TC values and DFR values, are also expected to contribute to the overall overprediction. As most of the used studies were combined with concurrent DFR data measurements, TC values could be derived.

In cereals (Table 4), one of the major field crops, the actual TC values

Table 3
Assumptions used to calculate exposure estimations.

Study	Default values	EFSA	Refined default values	Reference to refined default values
all	DFR (μg/cm^2 kg a.i./ha)	3	1.6	EUROPEOM II DFR_0 median value
	DT50 (Days)	30	10	EFSA (2008) which refers to Willis and McDowell (1987) "Bird and mammal risk"
	BW (kg)	60	70	EFSA Scientific Committee (2012)
a, c, n, o grapes	TC (cm^2/h)	10100	4600/3500/2651.76	BROV[a] 95th/75th percentile/mean
b, d, k, l, m, p pome and stone fruit	TC (cm^2/h)	4500	2815.85	EUROPOEM II[b] mean
e, f, g, h, I, j cereals	TC (cm^2/h)	1400	1054.36	ARTF studies[c] ARF 009 and 021 (mean)

NB: All TC values are for normal work clothing, but no gloves.

[a] BROV: HSE et al., 2020. Mean TC value for pruning/shoot-lifting; highest value from all tasks.

[b] EUROPOEM II. Mean TC value for hand-harvesting: van Hemmen et al. (2002) (EUROPOEM II project) proposed to consider a re-entry study conducted in peaches using azinphos-methyl (Spencer, 1991). Both potential and actual exposure was measured in the study using 10 replicates, and the mean transfer coefficient was derived by adding the mean potential hands transfer coefficient (1621.75 cm^2/h) and the mean potential body transfer coefficient reduced by a factor of 1° for protective clothing (1194.11 cm^2/h). The same approach and study was used to derive the currently used transfer coefficient for fruit trees, using the 90th percentile for hands and body), respectively.

[c] EFSA Guidance has used the 75th percentile values from the ARTF (Agricultural Re-entry Taskforce) studies ARF009 and ARF021 (Klonne, 1999a, 1999bbib_Klonne_1999abib_Klonne_1999b) for crop inspection for all crops (including cereals). The studies were carried out in dried peas and sweetcorn. The arithmetic mean value given here was taken from the combined datasets. The studies were reviewed by US EPA and used to derive a 1100 TC value (US EPA, 2017).

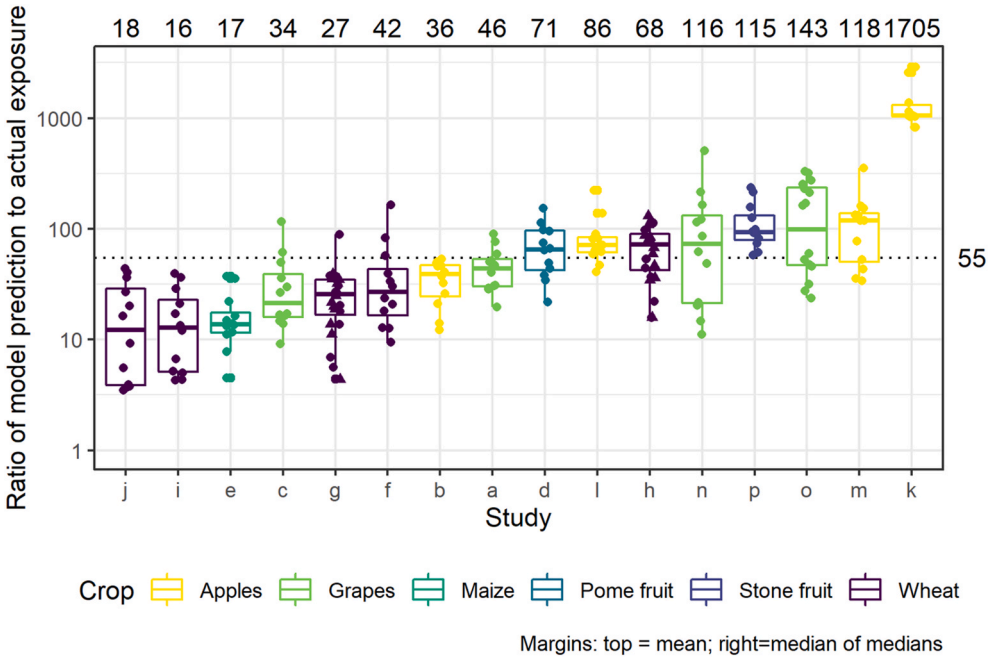

Fig. 3. Ratio of model predicted exposure estimate and actual external exposure by study of individual workers during re-entry, ordered by increasing study medians. Boxplots are generated over the individual worker exposure ratios, also if when two time points (point and triangle in studies g and h) were assessed.

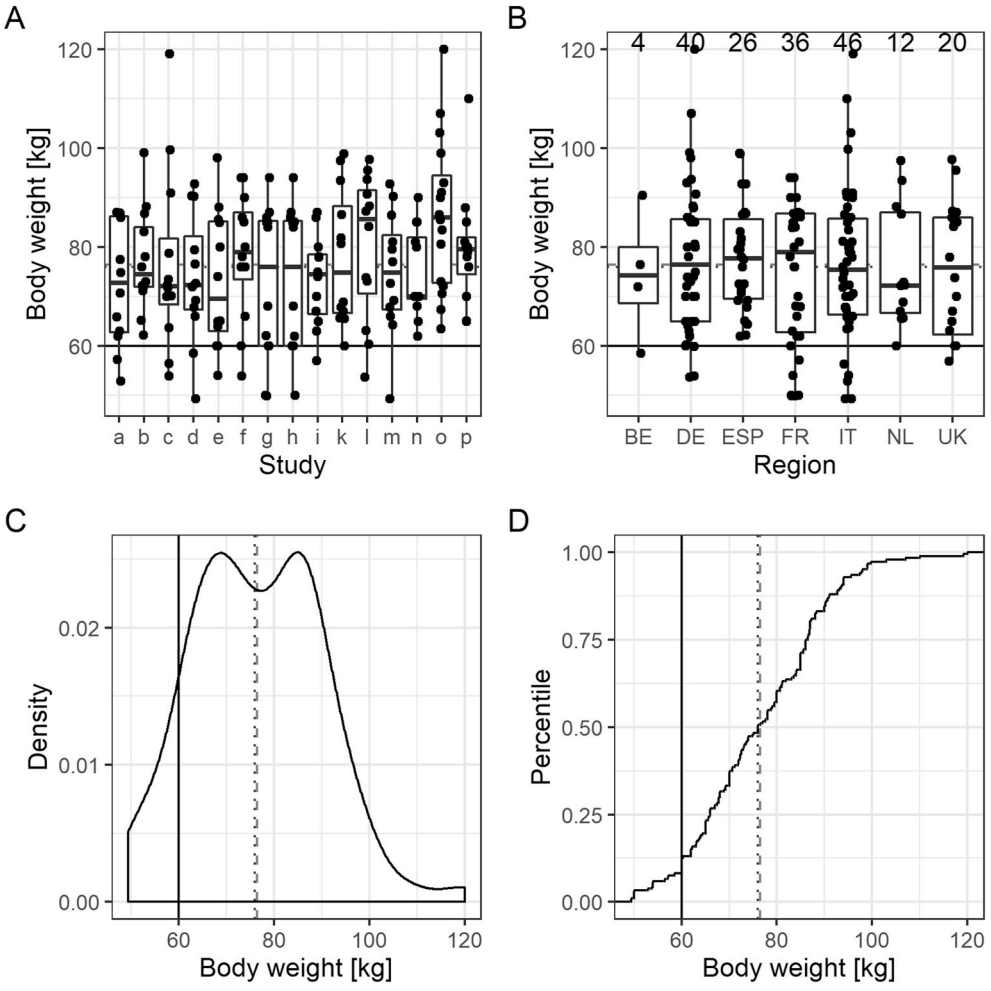

Fig. 4. Body weight distribution in the available worker studies indicate one reason for the overprediction seen in Fig. 3. Dashed red line indicates overall mean, dotted blue line (almost identical to the mean) overall median. A) Distribution per study. B) Distribution per country. C) Density distribution (all studies combined). D) Percentile of the distribution (all studies combined). The 60 kg assumed in the EFSA model systematically underpredicts actually measured worker body weight in the available studies. Note the bimodal distribution in Fig. 4C.

Table 4
TC values derived from combined worker exposure and DFR studies in cereals (in cm^2/h).

Study	Reference code	a.i.	Timepoint	Growth stage	Mean value	75th percentile
g	2015/1076589	1	0DALA	late	514	550
g	2015/1076589	1	1DALA	late	506	506
h	2015/1076589	2	0DALA	late	276	329
h	2015/1076589	2	1DALA	late	285	285
i	2014/1037490	1	0DALA	late	728	1161
i	2014/1037490	2	0DALA	late	770	1391
e	VV-470176	3	0DALA	early	201	277
f	VV-415297	4	0DALA	early	468	601
	Overall Mean				469	638

DALA = days after last application.

determined are all below the EFSA recommended TC of 1400 cm^2/h independent of whether early or late growth stages/crop heights were assessed. The mean value of all studies (469 cm^2/h) is by more than a factor of 2 lower to the current EFSA value. In study e (VV-470176) one worker received higher exposures than the rest. Although the exposure value obtained as such was not excluded, this value was obtained from a plot with an inexplicably very low DFR. By this combination the calculated TC value was an extreme outlier and was thus excluded. Therefore, TC values for this study come from 11 rather than 12 workers.

At times, conservatism in risk assessment can result from a lack of data directly relevant to the scenario in question. In the EFSA (2014) guidance, worker TC values for all scouting or crop inspection scenarios are derived from two US ARTF studies (ARF009 and ARF021). The first of these was carried out in sweetcorn where the crop was 1.83 m high. In ARF021, the crop was dried peas which at the final re-entry event were described as difficult to walk through. It is clear that in these studies, there will have been intensive contact with the treated crop. However, many pesticides are applied to arable crops where contact with the treated foliage is likely to be much less intense, e.g. early growth stage application of herbicides. Studies e and f included in this paper were carried out to provide more relevant data for such scenarios and returned 75th percentile TC values of 277 cm^2/h for maize at BBCH 14–18 (study e) and 601 cm^2/h for wheat at BBCH 25 (study f). These are clearly lower than the current EFSA default of 1400 cm^2/h. This provides an example of where a lack of data can (perhaps understandably) lead to overestimation of exposure. Although the best solution to this problem is to generate more specific data, it is unlikely every unique situation will be tested and there is a need for regulatory bodies to consider exactly what is involved in a particular scenario when deciding which datasets to extrapolate from, rather than simply applying the worst-case values. For pome fruit orchards (Table 5) only a very limited dataset is available but also these data indicate a substantial (>10-fold) overprediction by the EFSA guidance TC of 4500 cm^2/h, i.e. with a measured mean of 304 cm^2/h and a 75th percentile of 408 cm^2/h.

The underlying measured DFR values at the time of re-entry are also lower than what is used for the application rate calculated in the EFSA model.

It should be noted that a further relevant aspect of difference/overprediction of actual exposure is even not assessed by this comparison: the available studies based on passive dosimetry do not allow to determine systemic exposure. Thus, any contribution by calculatory use of dermal absorption estimates in relation to actual kinetics is not touched upon. The same applies for the included safety margins within the hazard assessment.

Table 5
TC values derived from combined worker exposure and DFR studies in apples (in cm^2/h).

Study	Reference code	a.i.	Timepoint	Growth stage	Mean value	75th percentile
k	2015/1113647	4	1DALA	Late, full foliage	138	157
m	2020/200620	5	1DALA	Late, full foliage	470	658
	Overall Mean				304	408

5. Alternative approaches

We described in Section 3 that the conservatism applied in hazard and exposure assessment provides many “hidden safety factors” for the re-entry risk assessment for workers, leading to compounded conservatism in risk assessment. We then demonstrated in Section 4 the extent of conservatism in one part of the risk assessment calculation, namely cumulative external exposure estimation for re-entry operations.

Generic overprediction prevents effective risk management and communication – the question is, how can we achieve better predictions? Four methods appear appropriate:

1) Transparently generate new and revise current individual default values in field or controlled settings and use in the current multiplication approach.
2) A generic exposure model based on measured data similar to predicting operator exposure in Europe.
3) Apply probabilistic methods on measured data.
4) Use mean point estimates instead of high percentiles in the current multiplication approach.

The approaches are shortly described and discussed in the following. It will become apparent that using mean point estimates in the current worker exposure estimation approach can be most easily and, more importantly, immediately implemented.

We also shortly describe how the conceptual exposure model could be refined.

5.1. *Transparently generate and revise default values*

If the current default values tremendously overpredict exposure, the immediate most obvious refinement strategy may be to revise the values based on novel data. The issue is that many studies addressing key exposure drivers are expensive and logistically challenging. It needs to be noted, that the selection of test items is currently driven by risk assessment model failure of individual products from individual companies, and data are only generated as higher tier refinements in product registration processes. Hence, one may discuss whether this selection is representative for the crop/use population to be addressed by the respective default values. There are currently no efforts to create pesticide exposure databases similar to efforts in the United States of America.

Further, for some default values more artificial and controllable experiments than the current field studies have not been discussed or developed, which would however allow the generation of more robust and specific exposure estimates. For example. the use of wind-tunnels is common when developing and testing spray nozzle performance and may be applicable for controlled exposure related studies as well. Furthermore, the field studies are always required to be with certain, often unrealistic worst case, meteorological conditions, which makes them harder to be designed and reproduced. Therefore, designing, conducting and interpreting studies in indoor/wind tunnel conditions is more desirable because of costs, resources and reproducibility.

Dermal absorption is one area where newly generated data based on

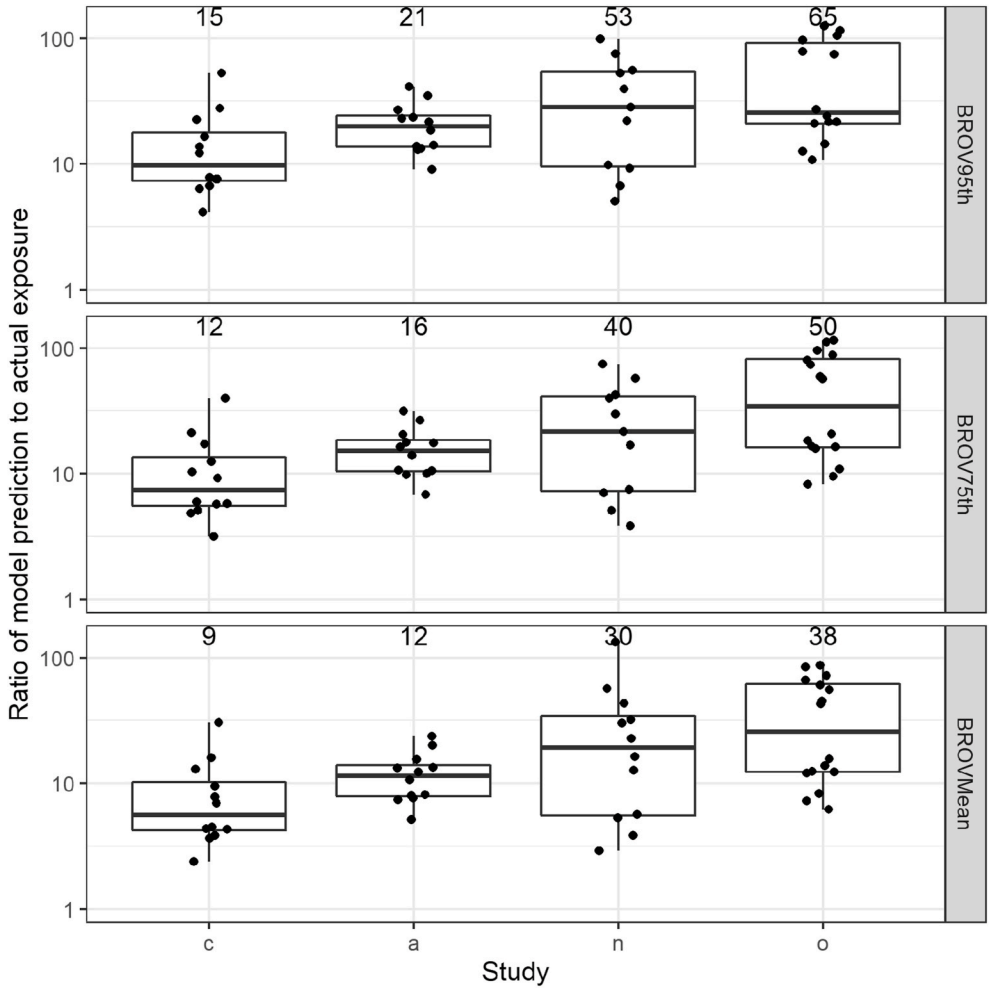

Fig. 5. Ratio of model predicted exposure estimate (using either the 95th, 75th percentile and mean) and actual external exposure by study of individual workers during re-entry (grapes). The mean ratio is shown at the top of the graphs.

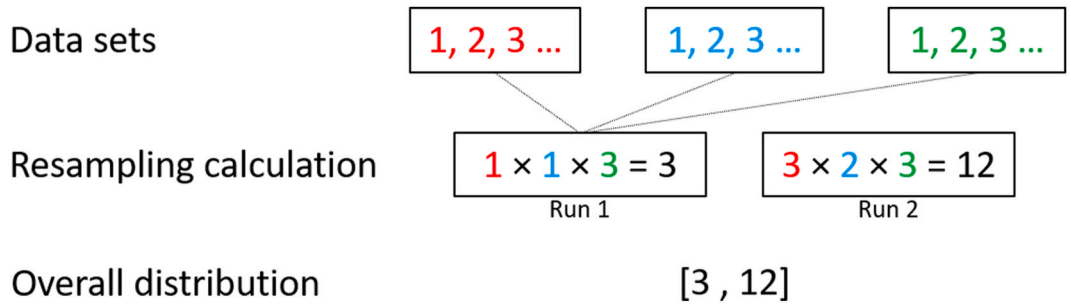

Fig. 6. Probabilistic exposure assessment schematic. Samples are taken repeatedly from the individual datasets to repeatedly calculate exposures. Based on the frequency of occurrence of the exposure drivers in the datasets, the overall exposure distribution will reflect the probability of potential exposures. This can be used directly or used to derive a single point value exposure estimate.

controlled experiments (Aggarwal et al., 2014, 2015bib_Aggarwal_et_al_2014bib_Aggarwal_et_al_2015) lead to revision of default values (EFSA, 2017b; EFSA, 2017c; EFSA Working Group on Dermal Absorption, 2015). As the studies showed differences between formulation types, more refined default values were proposed and derived for groups of product categories, different to those previously available (EFSA, 2012a). While the stakeholders disagreed on grouping, statistical derivation and the actual point values (Aggarwal et al., 2015; Chiusolo, 2017; Hothorn, 2017) and no consensus derivation approach was discussed, it at least shows that more information allows the derivation of more appropriate distinct values based on formulation properties.

Although the dermal absorption default values have been revised, the dermal absorption default values are still relatively high in practice and based on numerous worst-case assumptions. Hence, dossier submitters usually see a benefit in conducting product-specific data as a first measure of risk assessment refinement.

A similar straight-forward testing approach is not available for the external exposure estimations and relevant default values appear thus more important.

As described before, most of the default values have not been

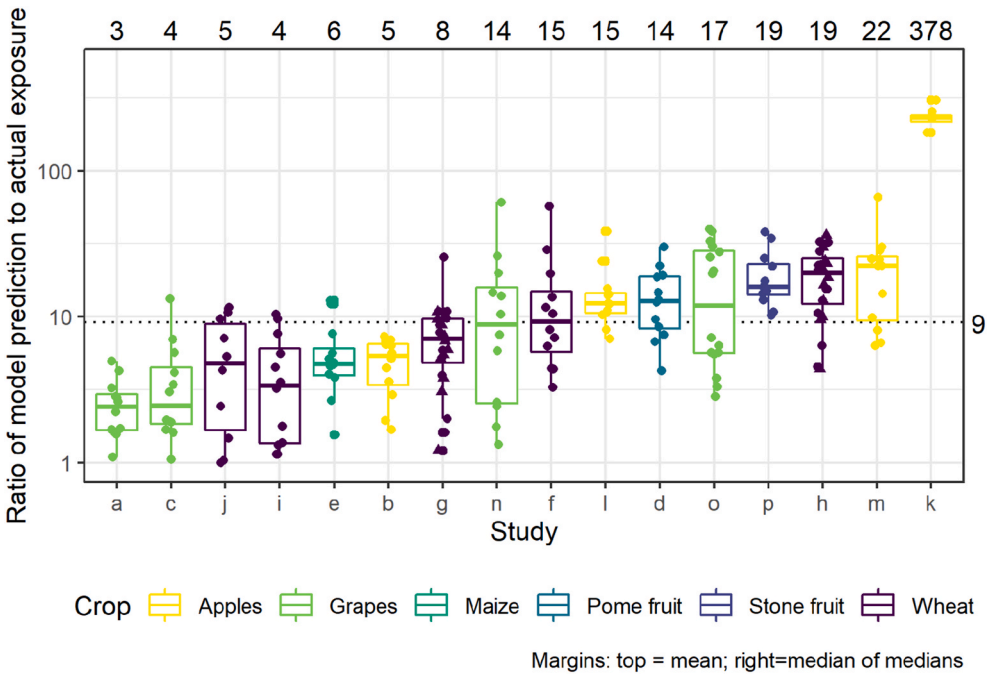

Fig. 7. This is similar to Fig. 3 and shows the ratio of model predicted exposure estimates and actual external exposure by study of individual workers during re-entry. In contrast to Fig. 3, however, average default values are used instead of high and non-harmonized percentile values. We see that the model prediction becomes much closer to the measured data, while still being conservative. The exposure of two workers (out of 184) is slightly underpredicted with a ratio of 0.924 and 0.981 (Study j).

transparently derived, their data base not published, or they are applied independent of formulation type or other properties. Also, the amount of data to derive default values is very heterogeneous. As an example, > 500 studies were used to derive dermal absorption default values while only a single study was used to derive a TC value for vegetables (harvesting and tying cucumbers in greenhouse).

Some initiatives to derive refined default values were however initiated, notably the BROV project (HSE et al., 2020) to derive more appropriate TC values for re-entry activities in vineyards. The BROV TC values achieve more meaningful exposure estimates compared to the EFSA model (Fig. 5). The figure also shows that using lower percentiles achieves lower overprediction, while still being protective.

5.2. Generic exposure model

In Europe, operator exposure is estimated with a regression-based generic calculation model, which includes results from 34 operator exposure studies, the Agricultural Operator Exposure Model (AOEM) (EFSA, 2014; Großkopf et al., 2013). There is unfortunately fewer data for worker re-entry exposure because those studies are even more expensive and logistically challenging than operator studies, and historically, workers were not considered to be the risk drivers of risk assessment of the various populations. Thus, there was no need to specifically address worker exposure at larger scale with experimental data.

An additional factor is that depending on the crop, the work scenarios are very different, for example with regard to harvesting activities, and less comparable than for example mechanical spray boom applications. However, it may obviously be explored whether the current dataset changes the *status quo*.

5.3. Probabilistic methods

Instead of using a generic model based on default values, which result in a single exposure estimate, it is possible to calculate a range of probability exposures based on resampling of the available datasets that were used to set default values. Here, random samples are repeatedly taken with replacement from the individual datasets, an exposure is calculated, and the overall distribution generated. The distribution relates to the probability for that exposure based on the data contained in the individual datasets. The process is shown in Fig. 6 and was previously conducted for the example calculation in Fig. 1. The benefit of the method is that it is not parametrically assuming a specific distribution and reflects the data most appropriately, including the probability of achieving certain net exposures, without losing information by deriving individual default values. The issues are that it might be considered to a certain extent as a black box, the individual data entries in the databases for the exposure drivers have to be available and the datasets need to be diligently curated for extreme values that bias the overall distribution (Lunchick, 2001). The effect of extreme values within the datasets can be mitigated by not using individual samples but means of >1 sample for the resampling calculation.

5.4. Using mean default values within the current approach

While the derivation of refined default values is presumably an approach that leads to better exposure estimates, there is still the statistical issue that the multiplication of multiple conservatively-derived default values will overpredict net exposure. The Agricultural Handler Exposure Task Force, L.L.C. (AHETF) and US EPA propose to use mean values for exposure estimates, because they consider that exposure realistically varies over time and regresses towards a mean (Crowley and Holden, 2019). Other benefits are that they are both easily understandable and directly applicable. Mean exposures are also used in the hazard characterization assays, when PODs (Point of Departure) are derived based on average exposure via food intake.

This approach is investigated in Fig. 7 using the default value assumptions shown in Table 3. Here, exposure estimations have been conducted as described above, however, instead of currently EU-wide accepted high percentile default values, averages have been used. The exposure estimates are much closer to the measured exposure, as compared to using high percentile values, while still being protective with an average overprediction of about 9-fold (based on the median of medians). Only two worker exposures (out of 184) were slightly underpredicted by the method (ratio of 0.924 and 0.981, both in study j).

While the multiplication approach might explain some of the remaining overprediction, impact of crop interception may be another driver not applied in the calculations.

5.5. Refining the conceptual exposure model

While using average default values to refine the external exposure estimation is the most pragmatic and straightforward approach, refining the conceptual exposure model is another possible method to derive more relevant doses for risk assessment, using information or data already available.

Relatively easy to implement is the adaption of the dose that re-entry workers are exposed to, see Fig. 2A. One could for example harmonize model assumptions between non-dietary exposure and environmental fate, by considering crop-interception factors (EFSA, 2017a).

Another approach is to revise the assumption regarding the internal dose. The default model assumes that the full exposure dose is available at t_0, while it can only be available after subsequent exposure events. Also, kinetic ADME parameters could be considered, which can be estimated by *in vitro* and *in vivo* surrogate values or by generic assumptions, see Fig. 8.

6. Discussion

Risk assessment needs to be conservative if there is uncertainty of the extent in which individual risk factors contribute to the overall risk. However, the propagation of conservatism, when it is applied on all individual risk factors independently, is currently not considered when developing European guidance documents. We showed the extent of what this means with respect to the simple example of estimating only the external worker re-entry exposure dose based on conservatively derived default values.

For informed risk management and realistic risk communication, it is desirable to have as precise exposure (and hazard) estimates as possible, because only then, risks can be efficiently weighed against each other. From a human health perspective, it is not desirable to lose an incorrectly failing (true) low-risk product and replace it with a (true) higher risk alternative. This type of Precautionary Principle-approach would lead to negative impact on the tools available to the farmer and agriculture in Europe.

We reviewed the existing conservatisms in European re-entry worker exposure and risk assessment (as an example), which offer much room for refinement. We demonstrated that the extent of model-driven risk overestimation can be refined by using different point estimates when setting default values. This modification achieves a more meaningful prediction of actual exposure based on the available data.

Here, only cumulative external worker re-entry exposure was estimated, which disregards other risk mitigation factors such as exposure kinetics, ADME and especially dermal absorption processes. Hence, the aggregated risk appears to be seriously overpredicted by the current European worker risk assessment approach.

If the amount of overprediction is unknown, it makes sense to use conservative estimates. However, when we can predict the cumulative external exposure better, e.g. by using average point estimates, we can discuss the level of conservatism that is intended when estimating exposure. The intended level of conservatism is currently not discussed by European stakeholders.

We show that using average point estimates for the current default values achieves a better, yet more reliably protective estimation of cumulative external exposure compared to using high and unharmonized percentile-based default values. This is also the approach for the Agricultural Handlers Exposure Task Force (AHETF) exposure data used by US EPA (Crowley and Holden, 2019).

Using averages has the benefit of accounting for variation over the assumed exposure duration. The risk assessment implicitly assumes subsequent daily exposures due to the used limit value, see Section 3, e.g. three months for a 90-day toxicity study. Note, this exposure duration might not be achievable for many scenarios such as harvesting of certain crops with a limited time window. Thus, default value averages assume that exposures regress towards an overall mean, with daily higher and lower exposures, which is reflected by the measured variation in the exposure studies and studies to refine the default values, e.g. the DFR.

Based on the available data, the average cumulative external worker re-entry exposure is overpredicted by about 9-fold when using average default values. Thus, it does not seem that an additional safety factor would be needed for a conservative assessment, when the current defaults are replaced with averages, since the exposure is almost certainly not underpredicted. The current point values for exposure estimation are not scientifically reasoned and are generated in isolation of each other. Hence, a refinement of these default values can be recommended based on the available data. Further, a harmonization between risk assessment assumptions of the different sections (toxicology, ecotoxicology and environmental fate) might also lead to more realistic risk estimates.

While the presented data allows a robust assessment of the conservatism associated with external exposure estimation, we may make an educated guess about other true but "hidden safety factors" within European worker re-entry risk assessment.

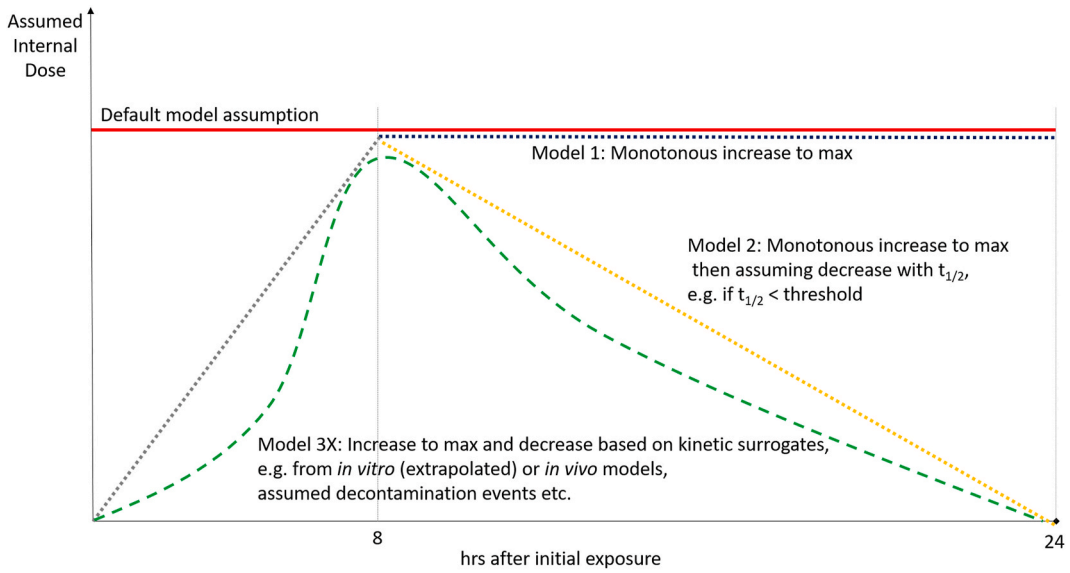

Fig. 8. Potential refinements of the internal dose.

We describe in Section 3 that kinetic considerations are not taken into account when estimating the internal dose, and that the use of cumulative dermal absorption and the assumption of receiving the full exposure dose at t = 0 of exposure probably results in a massive overprediction of the internal dose. We can speculate that this compounds, in the default model, to an overprediction of at least 100, assuming the factors in Sections 3.2 and 3.3 overestimate exposure about 25% on average, i.e. the ~20th root of 100.This would result in a joint factor of 5000 when considering the overprediction in the exposure model of about 50, for simplification. Alternatively, there is an approximate 500 000 margin of safety between a bioassay's NOAEL and the estimated dose a worker is exposed to during re-entry. While there is some biomonitoring data available for workers, there are only scarce data on true human NOAELs available for pesticides. If humans are similarly sensitive towards the adverse effects observed in animal studies, the margin of exposure will be even higher than 500 000, due to dose-spacing in the animal studies. However, such margins may be better explored with benchmark dose modelling to reduce the effect of dose-spacing on the point of departure (Haber et al., 2018; Hardy et al., 2017; Jensen et al., 2019). If humans are less sensitive towards adverse effects observed in animal studies, this factor obviously increases further. However, the potential "500 000 margin of safety" is obviously merely a guess, the presented data, however, allows a robust estimate of the actual exposure estimation for re-entry workers according to the current model.

Cochran and Ross (2017) previously reported the "unquantified" safety factor of being in the range of 47 to 1 000 000 for scenarios involving handlers, re-entry workers, and bystanders for US-based risk assessments of pesticides. Thus, our assessment for the European situation is in line with what they reported.

7. Conclusion

In conclusion, European worker re-entry risk assessment is very health-protective but may result in an unnecessary loss of products due to failed assessments based on the in-built conservatisms. Due to the many "hidden safety factors" within the method, such a loss seems unreasonable. Our data may thus be used to mitigate the scale of overprediction to some extent, as this can be easily reduced by using average default values, without any compromise to safety, and in turn, provides risk managers more options for refinements to provide the necessary tools for sustainable crop production.

Funding body information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The research was conducted as scientific expert contribution of the authors to the industry association, CropLife Europe.

CRediT authorship contribution statement

Felix M. Kluxen: Conceptualization, Resources, Visualization, Investigation, Supervision, Writing - original draft, Writing - review & editing. **Edgars Felkers:** Investigation, Formal analysis, Resources, Writing - original draft, Writing - review & editing. **Jenny Baumann:** Writing - original draft, Writing - review & editing. **Neil Morgan:** Resources, Data curation, Writing - review & editing. **Christiane Wiemann:** Resources, Data curation, Writing - original draft, Writing - review & editing. **Franz Stauber:** Writing - review & editing. **Christian Strupp:** Resources, Data curation, Writing - review & editing. **Sarah Adham:** Resources, Data curation, Writing - review & editing. **Christian J. Kuster:** Conceptualization, Writing - review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: All authors are employees of companies that conduct and evaluate risk assessments for regulatory purposes in the context of authorization and marketing of their companies' products. They contribute as scientific experts to the industry association CropLife Europe, former European Crop Protection Association, ECPA, for evaluation and development of the state-of-the-art methodology.

Acknowledgements

Christian J: Kuster (Chair), Edgars Felkers, Neil Morgan, Franz Stauber and Sarah Adham are members of the CLE Occupational and Bystander exposure expert team. Christiane Wiemann (Chair), Edgars Felkers and Felix M. Kluxen are members of the CLE Dermal absorption expert team. Christian Strupp (Chair) is a member of the CLE Human Health expert team. The authors would like to thank Steven McEuen (FMC) who organized the contact with the Agricultural Reentry Task Force, L.L.C. (ARTF) executive committee. We thank the ARTF, which allowed the use and publication of the mean TC value, based on ARTF studies, in this manuscript.

References

Aggarwal, M., et al., 2014. Assessment of in vitro human dermal absorption studies on pesticides to determine default values, opportunities for read-across and influence of dilution on absorption. Regul. Toxicol. Pharmacol. 68, 412–423.
Aggarwal, M., et al., 2015. Assessment of an extended dataset of in vitro human dermal absorption studies on pesticides to determine default values, opportunities for read-across and influence of dilution on absorption. Regul. Toxicol. Pharmacol. 72, 58–70.
Aggarwal, M., et al., 2019. Assessing in vitro dermal absorption of dry residues of agrochemical sprays using human SKIN within OECD TG 428. Regul. Toxicol. Pharmacol. 106, 55–67.
Barlow, S.M., et al., 2015. The role of hazard- and risk-based approaches in ensuring food safety. Trends Food Sci. Technol. 46, 176–188.
Blaschke, U., 2020. DFR meta analysis - an introduction to an ongoing ECPA/OBEEG project. In: 6th International Akademie Fresenius Conference "Worker, Operator, Bystander and Resident Exposure and Risk Assessment" +++ONLINE CONFERENCE +++.
Bormann, J.L., Maibach, H.I., 2020. Effects of anatomical location on in vivo percutaneous penetration in man. Cutan. Ocul. Toxicol. 1–10.
Brescia, S., 2020. Thresholds of adversity and their applicability to endocrine disrupting chemicals. Crit. Rev. Toxicol. 1–6.
Chiusolo, A., 2017. Changes to the Guidance on Dermal Absorption. Info Session on Applications – Pesticides - Technical Meeting with Stakeholders on EFSA Guidance on Dermal Absorption. EFSA.
Clarke, J.F., et al., 2018. Dermal absorption of pesticide residues. Chem. Res. Toxicol. 31, 1356–1363.
Cochran, R.C., Ross, J.H., 2017. A method for quantitative risk appraisal for pesticide risk assessments. J. Toxicol. Environ. Health 80, 1–17.
Cooper, A.B., et al., 2018. PBTK model for assessment of operator exposure to haloxyfop using human biomonitoring and toxicokinetic data. Regul. Toxicol. Pharmacol. 102, 1–12.
Creton, S., et al., 2010. Acute toxicity testing of chemicals-Opportunities to avoid redundant testing and use alternative approaches. Crit. Rev. Toxicol. 40, 50–83.
Crowley, M., Holden, L., 2019. Design and Implementation of Handler Monitoring Studies - Statistical Considerations. Joint Regulatory Technology Transfer Meeting; Session 3: Handler Scenarios (AHETF & ORETF).
Cullen, A.C., 1994. Measures of compounding conservatism in probabilistic risk assessment. Risk Anal. 14, 389–393.
Dankovic, D.A., et al., 2015. The scientific basis of uncertainty factors used in setting occupational exposure limits. J. Occup. Environ. Hyg. 12 (Suppl. 1), S55–S68.
Dourson, M.L., Stara, J.F., 1983. Regulatory history and experimental support of uncertainty (safety) factors. Regul. Toxicol. Pharmacol. 3, 224–238.
ECETOC, 2013. Technical Report No. 119: Evaluation of Systemic Health Effects Following Dermal Exposure to Chemicals. ECETOC, Brussels, Belgium.
EFSA, 2008. Risk assessment for birds and mammals - revision of guidance document under council directive 91/414/EEC (SANCO/4145/2000 – final of 25 September 2002) - scientific opinion of the panel on plant protection products and their residues (PPR) on the science behind the guidance document on risk assessment for birds and mammals. EFSA J. 6, 734.
EFSA, 2012a. Guidance on dermal absorption. EFSA J. 10, 2665.
EFSA, 2012b. Guidance on selected default values to be used by the EFSA Scientific Committee, Scientific Panels and Units in the absence of actual measured data. EFSA J. 10, 2579.
EFSA, 2014. Guidance on the assessment of exposure of operators, workers, residents and bystanders in risk assessment for plant protection products. EFSA J. 12, 3874.

EFSA, 2017a. EFSA Guidance Document for predicting environmental concentrations of active substances of plant protection products and transformation products of these active substances in soil. EFSA J. 15, e04982.
EFSA, 2017b. Guidance on dermal absorption. EFSA J. 15, 4873.
EFSA, 2017c. Info session on applications – pesticides - technical meeting with stakeholders on EFSA Guidance on dermal absorption. www.efsa.europa.eu.
EFSA Scientific Committee, 2012. Guidance on selected default values to be used by the EFSA Scientific Committee, Scientific Panels and Units in the absence of actual measured data. EFSA J. 10, 2579.
EFSA Working Group on Dermal Absorption, 2015. Assessment of new scientific studies on human in vitro dermal absorption. EFSA J. 13, 4304.
Elias, P.M., et al., 1981. Percutaneous transport in relation to stratum corneum structure and lipid composition. J. Invest. Dermatol. 76, 297–301.
European Comission, 2017. COMMISSION GUIDANCE DOCUMENT: Guidance on the Assessment of Exposure of Operators, Workers, Residents and Bystanders in Risk Assessment for Plant Protection Products SANTE-10832-2015 Rev. 1.7.
European Commission, 2008. REGULATION (EC) No 1272/2008 of the EUROPEAN PARLIAMENT and of the COUNCIL of 16 December 2008 on Classification, Labelling and Packaging of Substances and Mixtures, Amending and Repealing Directives 67/548/EEC and 1999/45/EC, and Amending Regulation (EC) No 1907/2006. Official Journal of the European Union. L 353/1.
European Commission, 2013a. COMMISSION REGULATION (EU) No 283/2013 of 1 March 2013 Setting Out the Data Requirements for Active Substances, in Accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council Concerning the Placing of Plant Protection Products on the Market. Official Journal of the European Union. L 93/1.
European Commission, 2013b. COMMISSION REGULATION (EU) No 284/2013 of 1 March 2013 Setting Out the Data Requirements for Plant Protection Products, in Accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council Concerning the Placing of Plant Protection Products on the Market.
European Commission SANCO, 2011. Guidance Document on the Preparation and Submission of Dossiers for Plant Protection Products According to the "Risk Envelope Approach.". EUROPEAN COMMISSION HEALTH AND CONSUMERS DIRECTORATE-GENERAL Safety of the Food Chain Chemicals contaminants, pesticides.
Feldmann, R.J., Maibach, H.I., 1967. Regional variation in percutaneous penetration of 14C cortisol in Man**From the division of dermatology. J. Invest. Dermatol. 48, 181–183. Department of Medicine, University of California School of Medicine, San Francisco, California 94122.
FOCUS, 2014. Generic Guidance for Tier 1 FOCUS Ground Water Assessments, Version 2.2.
Gelman, A., Greenland, S., 2019. Are confidence intervals better termed "uncertainty intervals"? BMJ 366, l5381.
Großkopf, C., et al., 2013. A new model for the prediction of agricultural operator exposure during professional application of plant protection products in outdoor crops. Journal für Verbraucherschutz und Lebensmittelsicherheit. 8, 143–153.
Gunther, C., et al., 2019. Comparison of in vitro and in vivo percutaneous absorption across human skin using BAY1003803 formulated as ointment and cream. Clin. Pharmacol. Drug Dev.
Guy, R.H., Maibach, H.I., 1984. Correction factors for determining body exposure from forearm percutaneous absorption data. J. Appl. Toxicol. 4, 26–28.
Guy, R.H., Maibach, H.I., 2002. Calculations of body exposure from percutaneous absorption data. In: Bronaugh, R.L., Maibach, H.I. (Eds.), Topical Absorption of Dermatological Products. Marcel Dekker, New York, NY, pp. 311–315.
Haber, L.T., et al., 2018. Benchmark dose (BMD) modeling: current practice, issues, and challenges. Crit. Rev. Toxicol. 48, 387–415.
Hardy, A., et al., 2017. Update: use of the benchmark dose approach in risk assessment. EFSA J. 15.
Hayes, A.W., 2014. Hayes' Principles and Methods of Toxicology. CRC Press, Taylor & Francis Group, Boca Raton.
Hothorn, L.A., 2017. Statistical Re-analysis to Derive Default Values. Info Session on Applications – Pesticides - Technical Meeting with Stakeholders on EFSA Guidance on Dermal Absorption. EFSA.
HSE, et al., 2020. Proposals for new transfer coefficient (TC) values for worker re-entry activities in vineyards - bystander Resident Orchard Vineyard (BROV) Re-entry Project Report. https://croplifeeurope.eu/media/news/bystander-resident-orchard-vineyard-brov-re-entry-project-report/.
Jackisch, R., et al., 2009. Inhibitory potency of choline esterase inhibitors on acetylcholine release and choline esterase activity in fresh Specimens of human and rat neocortex. J. Alzheim. Dis. : JAD. 16, 635–647.
Jensen, S.M., et al., 2019. A review of recent advances in benchmark dose methodology. Risk Anal. 39, 2295–2315.
Klonne, D.R., 1999a. Determination of Dermal and Inhalation Exposure to Reentry Workers during Scouting in Dry Peas. Unpublished. ARTF Study No. ARF021; Ricera, Inc. Document No. 7608-98-0111-CR-001; Grayson Research, LLC Study No. 98-326; MRID 45005908.
Klonne, D.R., 1999b. Determination of Dermal and Inhalation Exposure to Reentry Workers during Scouting in Sweet Corn. Unpublished. ARTF Study No. ARF009; H.E. R.A.C., Inc. Study No. 97-708HE; CAL Study No. 017-03; MRID 45005904.
Kluxen, F.M., 2019. Commentary: "New Statistics" in Regulatory Toxicology? https://doi.org/10.13140/RG.2.2.14639.48803 preprint available.
Kluxen, F.M., 2020. "New statistics" in regulatory toxicology. Regul. Toxicol. Pharmacol. 117, 104763.
Kluxen, F.M., et al., 2019. Dermal absorption study OECD TG 428 mass balance recommendations based on the EFSA database. Regul. Toxicol. Pharmacol. 108, 104475.
Kluxen, F.M., et al., 2020. Response to "OECD 428 in vitro dermal absorption mass balance performance based on our in-house database of pesticide studies. Regul. Toxicol. Pharmacol. 115, 104707.
Korpalski, S., et al., 2005. Dislodgeable foliar residues are lognormally distributed for agricultural re-entry studies. J. Expo. Sci. Environ. Epidemiol. 15, 160–163.
Krebs, B., et al., 2000. Uniform principles for safeguarding the health of workers re-entering crop growing areas after application of plant protection products. Nachrichtenblatt Dtsch. Pflanzenschutzd. 52, 5–9.
Latorre, A.O., et al., 2019. Non-relevance of acute dermal toxicity testing for assessing human health protection in the regulatory decision-making for agrochemical formulated products. Regul. Toxicol. Pharmacol.
Leese, H.J., et al., 2019. Going to extremes: the Goldilocks/Lagom principle and data distribution. BMJ Open 9, e027767.
Lehman, A.J., Fitzhugh, O.G., 1954. 100-fold margin of safety. Q. Bull. 18, 33–35. Association of Food & Drug Officials of the United States.
Lehman, P.A., et al., 2011. Percutaneous absorption in man: in vitro-in vivo correlation. Skin Pharmacol. Physiol. 24, 224–230.
Lev-Tov, H., Maibach, H.I., 2012. Regional variations in percutaneous absorption. J. Drugs Dermatol. JDD 11, e48–51.
Lunchick, C., 2001. Probabilistic exposure assessment of operator and residential non-dietary exposure. Ann. Occup. Hyg. 45, S29–S42.
Maibach, H.I., et al., 1971. Regional variation in percutaneous penetration in man. Arch. Environ. Health 23, 208–211.
McCarty, L.S., et al., 2020. Evaluation of the inherent toxicity concept in environmental toxicology and risk assessment. Environ. Toxicol. Chem.
Mielke, H., et al., 2017. Biometrical evaluation of the performance of the revised OECD Test Guideline 402 for assessing acute dermal toxicity. Regul. Toxicol. Pharmacol. 89, 26–39.
Morgan, N.A., et al., 2020. Dose setting for dermal absorption studies on dried foliar residues. Ann. Work Exposures Health. Accepted.
OECD, 2004. Test No. 428: Skin Absorption: in Vitro Method. OECD Publishing, Paris.
OECD, 2008. Test No. 407: Repeated Dose 28-day Oral Toxicity Study in Rodents. OECD Publishing, Paris.
OECD, 2011. GUIDANCE NOTES ON DERMAL ABSORPTION No. 156. OECD Publishing, Paris.
OECD, 2017. Test No. 402: Acute Dermal Toxicity. OECD Publishing, Paris.
PMRA, 2017. Acute Dermal Toxicity Study Waiver.
Ragas, A.M.J., 2011. Trends and challenges in risk assessment of environmental contaminants. J. Integr. Environ. Sci. 8, 195–218.
Ross, J.H., et al., 2000. Conservatism in pesticide exposure assessment. Regul. Toxicol. Pharmacol. 31, 53–58.
Rougier, A., et al., 1986. Regional variation in percutaneous absorption in man: measurement by the stripping method. Arch. Dermatol. Res. 278, 465–469.
Spencer, J.R., 1991. Long and short intervals of dermal exposure of peach harvesters to foliar azinphos-methyl residues. In: California. Division of Pest Management, E. P., Worker Safety. HS-1578. Worker Health and Safety Branch.
Strupp, C., et al., 2019. Phosmet: Growing Regulatory Uncertainty in Areas of Scientific Certainty (P5.1). IUPAC, Ghent.
US EPA, 2011. In: Assessment, N.C.f.E. (Ed.), Exposure Factors Handbook: 2011 Edition, EPA/600/R-09/052F. National Technical Information Servic, Washington, DC.
US EPA, 2012. Standard Operating Procedures for Residential Pesticide Exposure Assessment.
US EPA, 2016. Guidance for Waiving Acute Dermal Toxicity Tests for Pesticide Formulations & Supporting Retrospective Analysis.
US EPA, 2017. Science Advisory Council for Exposure (ExpoSAC) Policy 3.
van Beinum, W., Beulke, S., 2010. Collection and Evaluation of Relevant Information on Crop Interception for the Revision of the Guidance Document on Persistence in Soil, 7, p. 73E. EFSA Supporting Publications.
van Hemmen, J.J., et al., 2002. The Development, Maintenance and Dissemination of Generic European Databases and Predictive Exposure Models to Plant Protection Products: a EUROPOEM Operator Exposure Database: a EUROPOEM Bystander Exposure Database and Harmonised Model; a EUROPOEM Re-entry Exposure Database and Harmonised Model; an Evaluation of the Nature and Efficacy of Exposure Mitigation Methods; a Tiered Approach to Exposure and Risk Assessment, FAIR3 CT9vols. 6–1406. FAIR3 CT96-1406.
Vermeire, T., et al., 1999. Assessment factors for human health risk assessment: a discussion paper. Crit. Rev. Toxicol. 29, 439–490.
WHO, 1994. Environmental Health Criteria 170: Assessing Human Health RIsks of Chemicals: Derivation of Guidance Values for Health-Based Exposure Limits. WHO Document Production Services, Geneva, Switzerland.
WHO, 2005. Harmonization Project Document No. 2; Chemical-specific Adjustment Factors for Interspecies DIfferences and Human Variability: Guidance Document for Use of Data in Dose/Concentration-Response Assessment. WHO Document Production Services, Geneva, Switzerland.
Willis, G.H., McDowell, L.L., 1987. Pesticide persistence on foliage. In: W, G.W. (Ed.), Reviews of Environmental Contamination and Toxicology, vol. 100. Springer, New York, NY.
Yang, T., et al., 2017. Investigation of pesticide penetration and persistence on harvested and live basil leaves using surface-enhanced Raman Scattering mapping. J. Agric. Food Chem. 65, 3541–3550.
Zarn, J.A., et al., 2011. Study parameters influencing NOAEL and LOAEL in toxicity feeding studies for pesticides: exposure duration versus dose decrement, dose spacing, group size and chemical class. Regul. Toxicol. Pharmacol. 61, 243–250.