Chilton, ML, Api, AM, Foster, RS, Gerberick, GF, Lavelle, M, Macmillan, DS, Na, M, O'Brien, D, O'Leary-Steele, C, Patel, M, Ponting, DJ, Roberts, DW, Safford, RJ and Tennant, RE

Updating the Dermal Sensitisation Thresholds using an expanded dataset and an in silico expert system

Regulatory Toxicology and Pharmacology 133 (2022) 105200

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

Updating the Dermal Sensitisation Thresholds using an expanded dataset and an *in silico* expert system

Martyn L. Chilton [a,*], Anne Marie Api [b], Robert S. Foster [a], G. Frank Gerberick [c], Maura Lavelle [b], Donna S. Macmillan [a,1], Mihwa Na [b], Devin O'Brien [b,2], Catherine O'Leary-Steele [a], Mukesh Patel [a], David J. Ponting [a], David W. Roberts [d], Robert J. Safford [e], Rachael E. Tennant [a]

[a] *Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Leeds, LS11 5PS, UK*
[b] *Research Institute for Fragrance Materials, Inc., 50 Tice Boulevard, Woodcliff Lake, NJ, USA*
[c] *GF3 Consultancy, LLC, West Chester, OH, USA*
[d] *School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK*
[e] *B-Safe Toxicology Consulting, 31 Hayway, Rushden, Northants, NN10 6AG, UK*

ARTICLE INFO

Handling editor: Dr. Lesa Aylward

Keywords:
Derek Nexus
Dermal Sensitisation Threshold (DST)
Exposure-based waiving
In silico expert system
Quantitative Risk Assessment (QRA)
Skin sensitisation
Threshold of Toxicological Concern (TTC)

ABSTRACT

The Dermal Sensitisation Thresholds (DST) are Thresholds of Toxicological Concern, which can be used to justify exposure-based waiving when conducting a skin sensitisation risk assessment. This study aimed to update the published DST values by expanding the size of the Local Lymph Node Assay dataset upon which they are based, whilst assigning chemical reactivity using an *in silico* expert system (Derek Nexus). The potency values within the expanded dataset fitted a similar gamma distribution to that observed for the original dataset. Derek Nexus was used to classify the sensitisation activity of the 1152 chemicals in the expanded dataset and to predict which chemicals belonged to a High Potency Category (HPC). This two-step classification led to three updated thresholds: a non-reactive DST of 710 μg/cm^2 (based on 79 sensitisers), a reactive (non-HPC) DST of 73 μg/cm^2 (based on 331 sensitisers) and an HPC DST of 1.0 μg/cm^2 (based on 146 sensitisers). Despite the dataset containing twice as many sensitisers, these values are similar to the previously published thresholds, highlighting their robustness and increasing confidence in their use. By classifying reactivity *in silico* the updated DSTs can be applied within a skin sensitisation risk assessment in a reproducible, scalable and accessible manner.

1. Introduction

One of the many potential hazards that chemicals need to be assessed for is their ability to act as skin sensitisers, which can lead to allergic contact dermatitis in individuals exposed to sufficient levels of sensitising chemicals. Recent efforts in the field have sought to consolidate, strengthen, and formalise the procedure for assessing the risk of a chemical causing skin sensitisation through the Quantitative Risk Assessment (QRA)[3] framework (Api et al., 2008, 2020; Kimber et al., 2017). In this framework, an essential step is the use of the available skin sensitisation data (be that *in vivo*, *in vitro*, *in chemico*, *in silico*, and/or clinical) to derive a point of departure (often termed a No Expected Sensitisation Induction Level, or NESIL) measured in μg/cm^2, which is then used to calculate an Acceptable Exposure Level (AEL). If the expected Consumer Exposure Level (CEL) is less than the AEL, then it is considered that the exposure scenario(s) assessed will not lead to the induction of sensitisation. However, if the CEL is higher than the AEL, then the exposure scenarios are considered to pose an unacceptable risk of sensitisation.

One tool that was developed to help exposure-based risk assessments

Abbreviations: AEL, Acceptable Exposure Level; CEL, Consumer Exposure Level; DST, Dermal Sensitisation Threshold; EC3, Effective Concentration to cause a three-fold increase in lymphocyte proliferation; FN, False Negative; FP, False Positive; HPC, High Potency Category; HRIPT, Human Repeat Insult Patch Test; LLNA, Local Lymph Node Assay; NESIL, No Expected Sensitisation Induction Level; QRA, Quantitative Risk Assessment; S_NAr, Aromatic Nucleophilic Substitution; S_N2, bimolecular Nucleophilic Substitution; TN, True Negative; TP, True Positive; TTC, Threshold of Toxicological Concern.

* Corresponding author.
E-mail address: martyn.chilton@lhasalimited.org (M.L. Chilton).
[1] Present address: Humane Society International, 1255 23rd St NW, Washington, DC 20037, USA.
[2] Present address: Momentive Performance Materials, 769 Old Saw Mill River Road, Tarrytown, NY 10591, USA.

https://doi.org/10.1016/j.yrtph.2022.105200
Received 28 March 2022; Received in revised form 24 May 2022; Accepted 29 May 2022
Available online 1 June 2022

such as QRA is the Dermal Sensitisation Threshold (DST) (Safford, 2008; Safford et al., 2011, 2015), which is analogous to the Threshold of Toxicological Concern (TTC) used in the systemic toxicity risk assessment of chemicals (Cramer et al., 1976; Kroes et al., 2004; Munro et al., 1996). When exposure to a chemical is calculated to be lower than the relevant DST there is considered to be no appreciable risk of the chemical inducing skin sensitisation, as the DST represents a worst-case scenario based on the known skin sensitisation potency of hundreds of sensitisers. In these cases, an exposure-based waiving approach can be applied, thereby avoiding further data generation in the risk assessment.

Two separate DSTs were initially published in the literature: a value of 900 μg/cm^2 for non-reactive chemicals based on 38 compounds (Safford et al., 2011) and a value of 64 μg/cm^2 for reactive chemicals based on 233 chemicals (Safford et al., 2015). These were derived from the analysis of 363 chemicals which have been tested in the murine Local Lymph Node Assay (LLNA), which provides quantitative potency information in the form of an EC3 value, measured either as a percentage concentration (%) or as a dose per unit area (μg/cm^2). A set of structural rules was also published to identify any High Potency Category (HPC) chemicals, likely to have EC3 values below the reactive DST (Roberts et al., 2015). The proposed workflow for using the DST approach consists of answering two questions: is the chemical judged to be reactive using the published reaction mechanistic domains (Aptula et al., 2005), and if so, is it judged to be an HPC chemical? If the answer to the first question is no, then the non-reactive DST should be used; otherwise, if the answer to the second question is no, then the reactive DST should be used. Recently, a third DST of 1.5 μg/cm^2 was published, covering those chemicals that are predicted to belong to a High Potency Category (Nishijo et al., 2020).

One reason for wanting to update the Dermal Sensitisation Thresholds is that there is now a much larger amount of LLNA data available in the public domain than when the DST values were initially published. TTCs such as the DST are data-driven approaches, and therefore the more data used to derive them, the more accurate and robust the resulting values should be. The DST dataset used in the original publication has already been expanded once, with little change to the overall distribution of EC3 values (Safford et al., 2011). In this study, the dataset was expanded once again, and the reactivity classification step was automated and standardised using the *in silico* expert system Derek Nexus. Updated DST values were then derived from the expanded, classified dataset.

2. Material and methods

2.1. Compilation of the datasets

Information on the original DST dataset was taken from Safford et al. (2011). The dataset contains 363 chemical structures with an associated chemical name, CAS number (where available), expert-derived reaction mechanistic domain, EC3 value, and reference.

An expanded DST dataset was created by collecting and curating publicly available LLNA data. For positive LLNA outcomes, the EC3 value was recorded as a percentage concentration. For negative LLNA outcomes, the maximum tested dose was recorded, providing that the chemical had been tested to a concentration of 20% or greater. The chemical structures were standardised using a two-step procedure: firstly, they were normalised to ensure consistent representation of functional groups; and secondly, they were contextualised by removing stereochemistry, removing certain common counterions, and neutralising charges where possible. Any organometallic or inorganic structures that remained were removed from the dataset, as the DST approach is inappropriate for these chemicals (Safford et al., 2015). The resulting standardised structures were grouped on their InChI string (Heller et al., 2015) to ensure that each structure was only represented once in the final dataset. Where multiple positive LLNA outcomes had been observed for the same chemical, the median EC3 value was recorded to minimise the effect of any outliers in cases where a non-normal distribution of EC3 values was observed. Where multiple negative outcomes had been observed for the same chemical, the highest maximum tested dose was recorded. Chemicals with positive and negative LLNA outcomes recorded in the literature were conservatively assigned as sensitisers, and their median EC3 values were retained.

A small number of chemicals showing well-established false positive responses in the LLNA were identified and removed from the expanded DST dataset (n = 8, Table S1). The final expanded DST dataset contained 1152 chemicals (556 sensitisers and 596 non-sensitisers) with an associated CAS number (where available), EC3 value (or maximum tested dose in the LLNA), and reference.

2.2. In silico reactivity classification

The chemicals in the original and expanded DST datasets were classified into either the reactive or non-reactive domain using the skin sensitisation alerts found within Derek Nexus v6.1.1 (Lhasa Limited, 2021). These structural alerts describe toxicophores which are expected to cause sensitisation by combining an understanding of the mechanism of skin sensitisation with the available toxicity data and include a consideration of autoxidation and metabolism where these are relevant (for prehaptens and prohaptens, respectively). Predictions were processed with the following settings: knowledge base = Derek KB 2022 1.0, species = mammal, perceive tautomers = off, perceive mixtures = off, match alerts without rules = off, endpoint = skin sensitisation. The presence of one or more skin sensitisation alerts with an associated likelihood of either equivocal, plausible, or probable resulted in the chemical being classified as reactive. A prediction with a likelihood of improbable or a prediction of non-sensitiser based on a lack of structural alerts resulted in the chemical being classified as non-reactive.

The chemicals in the expanded DST dataset were classified as either HPC or non-HPC using the High Potency Category alerts found within Derek Nexus v6.1.1 (Lhasa Limited, 2021), using the same settings as above, except that endpoint = skin sensitisation HPC. Lipophilicity was calculated within Derek Nexus using BioByte's ClogP algorithm (version 5.9). The presence of one or more HPC alerts with an associated likelihood of plausible resulted in the chemical being classified as HPC. A lack of HPC alerts resulted in the chemical being classified as non-HPC.

Those chemicals classed as non-reactive by the skin sensitisation alerts in Derek Nexus, but HPC by the High Potency Category alerts, were conservatively classified as HPC. This reclassification occurred for 19 chemicals in the expanded DST dataset.

2.3. Data analysis

Dermal Sensitisation Thresholds were calculated by fitting a gamma distribution to the negative $\log_{10}$ transformed EC3 values as previously reported (Safford, 2008; Safford et al., 2011, 2015), and are reported to two significant figures. EC3 values were converted from % to μg/cm^2 by multiplying by a factor of 250 (Basketter et al., 2005). All data analysis was performed in KNIME v4.3 (Berthold et al., 2009), and the gamma distributions were fitted in R v3.6.2 using the fitdistr function within the MASS package. Parameters for the various gamma distributions used in this paper are available in Table S2.

The performance of the reactivity classifiers was assessed using the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions. From these various metrics were calculated, including accuracy ((TP + TN)/(TP + TN + FP + FN)), sensitivity (TP/(TP + FN)), specificity (TN/(TN + FP)) and balanced accuracy ((sensitivity + specificity)/2).

3. Results and discussion

3.1. Comparing in silico and human expert-derived reactivity classifications of the original DST dataset

The first question was how well an *in silico* expert system could classify chemicals as reactive or non-reactive compared to a human expert. The original DST dataset was used as a benchmark, and a direct comparison of the outcomes from the two reactivity classifiers showed that they agreed 86% (312/363) of the time. The performance of each was then compared against the LLNA data in the original DST dataset; Derek Nexus was found to have an almost identical sensitivity to human experts (87% cf. 86%) and very similar specificity (61% cf. 64%) (Table 1).

For the 243 chemicals which were classified as reactive by both human experts and Derek Nexus, a further comparison was made of which of the five mechanistic domains (Aptula et al., 2005) the chemicals were placed into: acyl transfer, Michael acceptor, Schiff base, S_N2, or S_NAr. The human expert classifications included the additional complex and special case categories, whereas Derek Nexus also contained categories for miscellaneous, hapten acting as a nucleophile, and free radical generator. The special case and miscellaneous categories were considered to be in partial agreement.

A total of 61% (148/243) of the chemicals were given an identical mechanistic domain by both classifiers, with a further 25% (60/243) showing partial agreement between the domains. In the remaining 14% (35/243) of cases, there was disagreement between Derek Nexus and human experts about which domain the chemical should be assigned. Since the DST approach is a binary classification of reactive/non-reactive, the identification of the precise mechanistic domain is not required for its use. However, the fact that Derek Nexus agrees or partially agrees with human experts on the mechanistic domain 86% of the time gives confidence that the skin sensitisation alerts in Derek Nexus have captured much of the same information as described by human experts in the published reaction mechanistic domains (Aptula et al., 2005).

Furthermore, as the performance metrics were almost identical (Table 1), Derek Nexus may be another valuable method for classifying skin sensitisation reactivity, alongside human expertise. There are several advantages to using an *in silico* expert system to classify chemical reactivity: 1) it automates and reduces the time taken to classify chemical reactivity, 2) the classifications are reproducible, standardised and accessible (as they can be made by organisations which lack the relevant in house sensitisation expertise to manually classify chemicals), and 3) detailed alert comments and relevant references support the reactive outcomes.

Table 1
Comparison of the performance of both human and *in silico* reactivity classifiers against the original and expanded DST datasets.

Reactivity classifier	Dataset	Accuracy	Sensitivity	Specificity	Balanced accuracy
Human expert	Original	80.4%	86.0%	64.1%	75.1%
Derek Nexus	Original	80.2%	86.7%	60.9%	73.8%
Derek Nexus	Expanded	73.6%	85.4%	62.6%	74.0%

3.2. Using in silico reactivity classifications to recalculate the original Dermal Sensitisation Thresholds

The *in silico* reactivity classifications generated by Derek Nexus were used to recalculate the reactive and non-reactive DSTs from the original DST dataset to see what effect changing from the human expert reactivity classifier would have. The 235 reactive chemicals (as classified by Derek Nexus) were used to derive a reactive DST to compare against the initially published value of 64 μg/cm^2. Firstly, a gamma distribution was fitted to the reactive sensitisers in the dataset (Fig. S1). Secondly, using Derek Nexus's calculated sensitivity and specificity values of 87% and 61%, coupled with the estimated global incidence of sensitisers of 20% (Safford, 2008), the probability of a reactive chemical being a true sensitiser (i.e. the positive predictive value) was calculated to be 35.7% (Equation (1)). Given that the reactive DST is designed to leave only 5% of chemicals likely to have a potency higher than the threshold, this means that the probability of a reactive chemical being a less potent sensitiser than the DST is equal to (1 – (0.05/0.357)) = 0.860. The EC3 value associated with the 86th percentile of the gamma distribution was 77 μg/cm^2 (Fig. S1).

$$\textit{Positive predictive value} = \frac{\textit{sensitivity} \times \textit{prevalence}}{(\textit{sensitivity} \times \textit{prevalence}) + ([1 - \textit{specificity}] \times [1 - \textit{prevalence}])}$$

$$= \frac{0.867 \times 0.2}{(0.867 \times 0.2) + ([1 - 0.609] \times [1 - 0.2])} = \frac{0.173}{0.173 + 0.313} = 0.357$$

Equation (1). Calculation of the probability that a reactive chemical within the original DST dataset is a sensitiser.

A similar process was employed to calculate a comparative non-reactive DST using the 36 non-reactive sensitisers (as classified by Derek Nexus) in the original DST dataset, and a gamma distribution was fitted to the EC3 values (Fig. S1). In accordance with the published methodology (Safford et al., 2011), the 95th percentile of this distribution was taken as the non-reactive DST, which gave a value of 580 μg/cm^2.

The reactive DST based on reactivity assignments from Derek Nexus only differed by 1.2-fold from the originally published value (77 cf. 64 μg/cm^2). This demonstrates that the classifier used (Derek Nexus or human expert) does not create significant differences in either the gamma distribution fitted to the EC3 values of the reactive sensitisers, or the DST value derived from this distribution (Fig. S1). However, the non-reactive DST showed a slightly larger deviation of 1.6-fold from the published value (580 cf. 900 μg/cm^2). Visual inspection of the gamma distribution fitted to the EC3 values of the non-reactive sensitisers shows that this is primarily due to the presence of one chemical (hexyl salicylate) classified as non-reactive by Derek Nexus and only marginally reactive by human experts (Table S1). Hexyl salicylate gives an unusually strong response in the LLNA (EC3 = 45 μg/cm^2) when compared to its (at worst) weakly sensitising results in humans. The no observed effect level in the Human Repeat Insult Patch Test (HRIPT) was the maximum tested dose of 35,433 μg/cm^2 (Api et al., 2015), which is approximately 800-fold less potent than the LLNA outcome. Based on an analysis of the available human sensitisation data for hexyl salicylate, it was assigned to human potency category 4, defined as "substances that are rarely important clinical allergens, because they require considerable/prolonged exposure to higher dose levels to produce sensitisation, which even then is unlikely to exceed 0.01% of those exposed" (Basketter et al., 2014). The discrepancy between the murine and human

data has been rationalised by suggesting that hexyl salicylate is a surface-active irritant (Roberts and Api, 2018), which would explain the apparent over-prediction of potency by the LLNA. Case reports describing allergic contact dermatitis to other salicylates are available, but rare (Miralles et al., 2015; Mortz et al., 2010; Shaw, 2006; Singh and Beck, 2007). The fact that Derek Nexus no longer predicts salicylates as sensitising is reflective of the evidence outlined above, alongside other available literature data for this class of chemicals (Montelius et al., 1998; Spielmann et al., 2007). The disproportionate impact of this single chemical on the non-reactive DST can be seen by removing this chemical from the distribution; the resulting non-reactive DST derived from the remaining 35 non-reactive sensitisers would be 930 μg/cm^2, which is almost identical to the previously published value of 900 μg/cm^2. The hexyl salicylate example demonstrates the more general point that the non-reactive DST is more sensitive to potent outliers than the reactive DST, as these can have a large impact on the location of the 95th percentile within the tail of the gamma distribution.

3.3. Comparing the original and expanded DST datasets

Having established that Derek Nexus could be reliably used to classify skin sensitisation reactivity, the two DST datasets were compared. The expanded dataset contained over three times as many chemicals as the original dataset (n = 1152 cf. n = 363), and the incidence of sensitisers within the two also differed significantly, with the original dataset skewed towards sensitisers (75:25) while the expanded dataset was approximately balanced (48:52). Consequently, the expanded dataset contained twice as many sensitisers as the original dataset (n = 556 cf. n = 271). Despite this, the overall EC3 distribution was comparable to that derived from the original dataset (Fig. 1) (Safford et al., 2011). This comparison corroborates earlier findings that the distribution of known sensitisers is both stable and well-modelled using a gamma distribution (Safford, 2008; Safford et al., 2011).

An analysis was conducted to compare the number of sensitisers falling into each mechanistic domain for each dataset, to highlight areas where additional data is now available (Table S3). Within the original dataset, 32% of the sensitisers were Michael acceptors (86/271), 18% were S_N2-reactive (48/271), 16% Schiff base formers (43/271), 10% acyl transfer agents (28/271) and 1% S_NAr-reactive (4/271). The remaining sensitisers were split between those classed as non-reactive by a human expert (n = 38), those having multiple domains (n = 3) and complex or special cases (n = 21). Within the expanded dataset, 24% of the sensitisers were Michael acceptors (133/556), 21% were Schiff base formers (114/556), 20% S_N2-reactive (110/556), 9% acyl transfer agents (52/556) and 4% S_NAr-reactive (21/556). The remaining sensitisers were split between those classed as non-reactive by a human expert (n = 76), those having multiple domains (n = 31), nucleophilic sensitisers (n = 15) and complex or special cases (n = 4). Although the distribution of the five main mechanistic domains was broadly similar across the two datasets, the domains which saw the largest proportional increase in the number of sensitisers were the S_NAr-reactive (5.3-fold increase), Schiff base (2.7-fold increase) and S_N2-reactive (2.3-fold increase) domains.

The performance of Derek Nexus in classifying the chemicals as reactive or non-reactive was compared for the two datasets (Table 1). The sensitivity remained essentially constant (85% cf. 87%) when increasing the size of the dataset from 363 to 1152 chemicals, as did the specificity (63% cf. 61%). One difference between the two datasets is the incidence of sensitisers within them, which has decreased from 75% to 48%. While neither of these figures is representative of the true incidence of sensitisers in the known chemical universe (estimated at 20%, see Safford, 2008), when analysing binary classifiers, it is better to have a balanced dataset where possible, as this gives a more representative picture of the system's performance (Wei and Dunbrack, 2013).

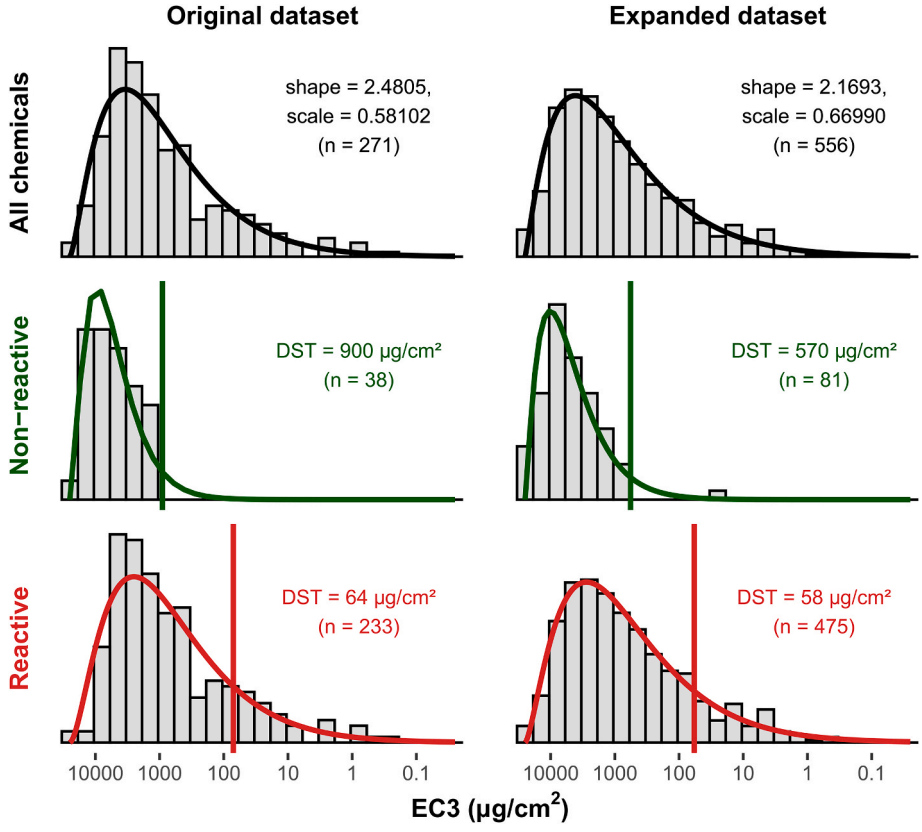

Fig. 1. Comparison of the gamma distributions and corresponding DSTs for the sensitisers in the expanded DST dataset with those published for the original DST dataset.

A gamma distribution was fitted to the negative logarithm of the EC3 values of the 475 reactive sensitisers in the expanded DST dataset (Fig. 1). By applying Equation (1) and using Derek Nexus's calculated sensitivity and specificity values of 85% and 63%, the probability of a reactive chemical being a true sensitiser was calculated to be 36.3%. Given that by design, only 5% of chemicals are expected to have an EC3 value below the reactive DST, the probability of a reactive chemical being a less potent sensitiser than the DST was calculated to be (1 – (0.05/0.363)) = 0.862. The EC3 value associated with the 86.2nd percentile of this gamma distribution was 58 μg/cm^2. This value is only 1.1-fold smaller than the previously calculated reactive DST of 64 μg/cm^2. Considering this is based on twice the number of chemicals found in the original dataset (n = 475 cf. n = 233), the fact that the resulting value is so close to the original value demonstrates the remarkable robustness of this threshold.

A gamma distribution was also fitted to the negative logarithm of the EC3 values of the 81 non-reactive sensitisers in the expanded DST dataset (Fig. 1). The EC3 value associated with the 95th percentile of this gamma distribution was 570 μg/cm^2. This value is 1.6-fold smaller than the previously calculated non-reactive DST of 900 μg/cm^2. If the eight well-established false positive chemicals were not excluded from the expanded dataset (Table S1), the magnitude of the resulting non-reactive DST would have been similar (510 μg/cm^2, calculation not shown). This information demonstrates that the decision to remove these chemicals from the dataset due to concerns about the quality of their biological data did not significantly impact the results of the research. However, the inclusion of the false positive hexyl salicylate would cause a slight reduction in the resulting DST.

Table 2
Summary of minor updates to the HPC rules, based on analysis of the expanded DST dataset.

HPC rule	Summary of changes
1 – Protein derivatisation agents	Addition of carbodiimides. Addition of isothiazolinones. Addition of ortho-phthalaldehydes and other 1,4- and 1,5-dialdehyde and ketoaldehyde cross-linkers.
2b – Direct acting Michael acceptors (with a single activating group)	Exclusion of activated alkenes which do not have a beta-hydrogen.
2c – Direct acting Michael acceptors (with more than one activating group)	Addition of acetylenic aldehydes, ketones, esters and amides. Exclusion of activated alkenes which do not have a beta-hydrogen (unless in a strained ring).
3a – Pro/pre-Michael acceptors	Addition of aminopyrroles and aminopyrazines. Addition of 1,5-, 1,7-, and 2,6-diamino, dihydroxy and aminohydroxy naphthalenes and azo analogues. Specification that diaminobenzenes must have at least one primary/secondary amine.
4a – Schiff base electrophiles	Specification of distinct logP limits for different classes of 1,2-dicarbonyl.
6a – S_N2 electrophiles (benzylic halides and pseudohalides)	Exclusion of alpha-halo carboxylic acids.
6b – S_N2 electrophiles (methylating agents)	Addition of primary N-alkyl nitroso amido compounds.
6c – S_N2 electrophiles (special cases)	Addition of nitrogen and sulphur mustards. Addition of suitably lipophilic activated epoxides.
7a – S_NAr electrophiles (with more than one activating group)	Addition of 5-membered aromatic rings with two ortho activating groups (namely groups of equal or greater electronegativity than a nitro group and/or ring nitrogen atoms). Removal of thiols, aldehydes, ketones and substituted hydrocarbons from the list of allowed leaving groups.

3.4. Updating the HPC rules

As part of the exercise of comparing the original and expanded DST datasets, an analysis of the originally published High Potency Category rules (Roberts et al., 2015) was conducted to check that they could still identify those sensitisers with a potency expected to be below the reactive DST. The 556 sensitisers in the expanded DST dataset were manually assessed for reactivity and subsequently whether they should be classified as HPC, based on their chemical structure and an understanding of their reactivity and (when required) lipophilicity. This analysis revealed several toxicophores in the expanded dataset which were not present in the original dataset, but which should be considered alerts for HPC chemicals (Table 2). As well as the inclusion of these new toxicophores in the updated HPC rule set, a few of the rules were amended by restricting the scope to remove a small number of substituents that are no longer believed to cause a chemical to be HPC. This highlights the synergy that can be achieved between human and *in silico* expert knowledge and shows the importance of updating (*in silico* and human) expert knowledge as new data become available. The rule changes are described in further detail in the Supplementary Information.

The resulting HPC rules were then encoded into Derek Nexus as structural alerts to enable the HPC classification also to be automated. The encoded HPC rules achieved the same classification as that of a human expert for 92% (511/556) of the structures. The most common reason for a difference in classification (n = 25) was that a chemical was classified as HPC by a human expert according to rule 9 (structural complexity), which, due to the subjective nature of this rule (representing uncertainty about the reactivity of an unusual functional group), cannot be easily encoded within *in silico* expert system.

As expected, the distribution of those chemicals classified as HPC was heavily skewed towards the more potent sensitisers, and most sensitisers with an EC3 value below the previously published reactive DST (64 μg/cm^2) were classified as HPC (59/69, 86%). Almost every chemical identified as HPC (144/146, 99%) was also classified as reactive by Derek Nexus, although two chemicals were classified as HPC despite a lack of skin sensitisation alerts for reactivity. Therefore, a conservative approach was applied, and these chemicals were removed from the non-reactive domain and placed into the HPC domain for all future analysis.

The analysis and subsequent modifications to the HPC rules show that these rules continue to be conservative and are almost always protective. The small number of potent sensitisers (n = 10) which were not classified as HPC are discussed in due course (Table 3).

3.5. Using the expanded DST dataset and updated HPC rules to calculate updated Dermal Sensitisation Thresholds

The HPC rules were created as a belt-and-braces approach to filter out very potent chemicals from the reactive domain after deriving the reactive DST, to ensure that it was suitably protective. In this study another opportunity to use them was identified, namely to separate the HPC chemicals from the non-HPC chemicals in the reactive domain prior to deriving the reactive DST, thus creating the following reactivity classifications: non-reactive, reactive (non-HPC), and HPC. From this position, three updated DSTs could be derived from the expanded dataset: a non-reactive DST based on those chemicals classified as non-reactive according to the skin sensitisation alerts in Derek Nexus; a reactive (non-HPC) DST based on those chemicals classified as reactive but not classified as HPC according to the High Potency Category rules in Derek Nexus; and an HPC DST, based on those chemicals classified as HPC.

Classifying the 556 sensitisers in the expanded DST dataset using this approach led to 79 sensitisers falling into the non-reactive domain, 331 sensitisers falling into the reactive (non-HPC) domain, and 146 sensitisers falling into the HPC domain. A gamma curve was fitted to each of

Table 3
The 16 sensitisers in the expanded dataset which have an EC3 value lower than the corresponding DST.

Name	CAS number	Median EC3 (μg/cm^2)	Class	Comments
Non-reactive sensitisers with an EC3 < 710 μg/cm^2				
2,7-Naphthalenediol	582-17-2	700	Misclassification	Classified as reactive by human expert judgement. The reaction chemistry of this compound is likely to be similar to that of resorcinol, where one OH group activates the carbon atom ortho to the other OH group, making the ortho carbon doubly activated as a nucleophilic and/or free radical reaction site.
1,1'-(1,1,2,2-Tetramethyl-1,2-ethanediyl) bisbenzene	1889-67-4	675	FP – impurity	Likely to contain potent impurities.
(1R,4S,4aR,8aS)-9-(Dichloromethylidene) octahydro-1,4-methanonaphthalene-5,8-dione	1369500-14-0	650	FP – impurity	Likely to contain potent impurities (e.g., retro-Diels-Alder reactions forming benzoquinone or hydroquinone).
Reactive (non-HPC) sensitisers with an EC3 < 73 μg/cm^2				
2,6-Diaminopyridine	141-86-6	62.5	Complex	Possible nucleophilic sensitiser, but complex chemistry is not obvious from inspection of the structure.
3-Hydroxy-N-(2-methoxyphenyl)-2-naphthamide	135-62-6	50	FP – irritant	Structure suggests similar physical chemistry to hexyl salicylate, which is considered an LLNA FP with an EC3 below the reactive DST (Roberts and Api, 2018).
N-(Oxiran-2-ylmethyl)aniline	No CAS	47.5	Misclassification	Classified as HPC by human expert judgement using rule 9, due to structural complexity. Oxidation of the NH–CH_2 group would give glycidaldehyde which is predicted to be much more reactive (Roberts et al., 2017).
Chlorpromazine, Chlorpromazine hydrochloride	50-53-3, 69-09-0	35	Misclassification	Classified as HPC by human expert judgement using rule 9, due to structural complexity. A rare case of a pro-Schiff base aliphatic amine being more potent than the reactive DST, likely because of the high lipophilicity of the prohapten, and the specific reaction chemistry of the resulting hapten malondialdehyde (Slatter et al., 2000).
Trimethylolpropane triacrylate	15625-89-5	17.71	Misclassification	Classified as HPC by human expert judgement using rule 2b (MA with a single activating group). While a single reactive sub-structure CH_2=CH–C(=O)- is not alone sufficient to trigger HPC classification, the presence of more than one such substructure is expected to enhance the potency due to their cross-linking ability.
2-(3,4-Dimethylphenyl)-5-methyl-1,2-dihydro-3H-pyrazol-3-one	18048-64-1	14.25	Over-estimation of potency	The potency (NICEATM, 2010) may be over-estimated due to extrapolation from the dose-response data (Ryan et al., 2007). Potency estimated to be ~330 μg/cm^2 using the probit extrapolation method (Roberts, 2015). For comparison, the very similar analogue 3-methyl-1-phenylpyrazolone has an EC3 value derived by linear interpolation which is two orders of magnitude higher, equal to 2125 μg/cm^2 (Kern et al., 2010).
1-Benzoylacetone	93-91-4	10	Unreliable LLNA data	EC3 value seems extraordinarily low, and unfortunately the dose-response data is not available for interrogation (Ashikaga et al., 2010). For comparison, six very similar 1,3-diketone analogues have EC3 values three orders of magnitude higher, in the range 2075–7250 μg/cm^2 (Gerberick et al., 2005).
3,3′,4′,5-Tetrachlorosalicylanilide	1154-59-2	10	Misclassification	Classified as HPC by human expert judgement using rule 9, due to structural complexity (Roberts et al., 2015).
N-[(1R,2E,4S,6S)-1-methyl-4-(prop-1-en-2-yl)-7-oxabicyclo[4.1.0]heptan-2-ylidene] hydroxylamine, N-[(1S,2E,4S,6R)-1-methyl-4-(prop-1-en-2-yl)-7-oxabicyclo[4.1.0]heptan-2-ylidene] hydroxylamine	No CAS	4.5	Misclassification	Classified as HPC by human expert judgement using rule 9, due to structural complexity. Able to tautomerise to a nitroso-olefin, which would be a highly reactive Michael acceptor (Bergström et al., 2008).
(4R,5S)-1,5-Dimethyl-3-(1-oxo-2-propenyl)-4-phenyl-2-imidazolidinone	139109-23-2	1.1	Misclassification	Classified as HPC by human expert judgement using rule 9, due to structural complexity. Uncertainty about the electronegativity of the –C(=O)NHC=O activating group means that reactivity relative to known Michael acceptor sensitisers cannot be confidently predicted.
HPC sensitisers with an EC3 < 1.0 μg/cm^2				
4′-Hydroxychalcone	2657-25-2	0.40	Over-estimation of potency	The potency (NICEATM, 2010) may be over-estimated due to extrapolation from the dose-response data (Ryan et al., 2007). Potency estimated to be ~45 μg/cm^2 using the probit extrapolation method (Nishijo et al., 2020; Roberts, 2015).
4-Ethoxymethylene-2-phenyl-2-oxazolin-5-one	15646-46-5	0.35	Very potent	Exceptionally potent sensitiser (Loveless et al., 1996; NICEATM, 2010), suggested to be due to its particularly fast and diverse reaction chemistry (Natsch et al., 2010).
Benzo[a]pyrene	50-32-8	0.225	Over-estimation of potency	The potency (Gerberick et al., 2005) may be over-estimated due to extrapolation from the dose-response data (Ryan et al., 2007). Potency estimated to be ~6 μg/cm^2 using the probit extrapolation method (Roberts, 2015).

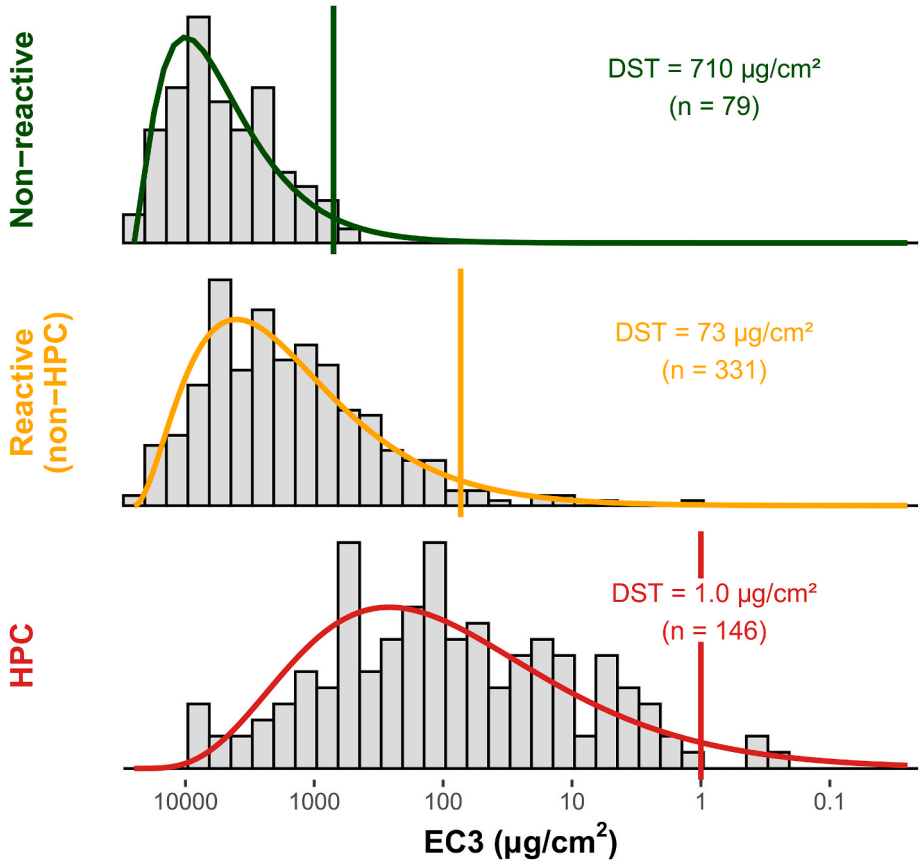

Fig. 2. Gamma distributions and updated DSTs for non-reactive, reactive (non-HPC), and HPC sensitisers based on the expanded DST dataset.

these distributions to derive DSTs for each class of chemicals (Fig. 2).

The updated non-reactive DST associated with the 95th percentile of the gamma distribution of the 79 non-reactive sensitisers was 710 µg/cm^2. This value is similar to, and only 1.3-fold smaller than, the previously published non-reactive DST of 900 µg/cm^2. Using the performance figures for the subset of the expanded dataset that was classified as non-HPC (sensitivity = 81%, specificity = 65%), the probability that an unknown chemical which is classed as non-reactive would be a sensitiser (i.e. 1 – negative predictive value) can be calculated to be 6.9% (Equation (2)). The probability of a non-reactive chemical having an experimental EC3 value less than 710 µg/cm^2 is therefore 0.05 × 0.069 = 0.0035, meaning that the updated non-reactive DST is 99.7% protective (i.e., there is a 99.7% probability that a chemical classed as non-reactive will either be non-sensitising or have an EC3 value greater than 710 µg/cm^2).

$$1 - \textit{negative predictive value} = \frac{\textit{specificity} \times (1 - \textit{prevalence})}{([1 - \textit{sensitivity}] \times \textit{prevalence}) + (\textit{specificity} \times [1 - \textit{prevalence}])}$$

$$= 1 - \frac{0.645 \times (1 - 0.2)}{([1 - 0.807] \times 0.2) + (0.645 \times [1 - 0.2])} = 1 - \frac{0.516}{0.039 + 0.516} = 1 - 0.931 = 0.069$$

Equation (2). Calculation of the probability that a non-reactive chemical within the expanded DST dataset is a sensitiser.

Similarly, the 95th percentile was taken from the gamma distribution of the 331 reactive (non-HPC) sensitisers. While the previously published reactive DST used the 86th percentile (calculated so that the overall DST would be 95% protective before the HPC rules were applied), the approach described above meant that a different percentile was now appropriate. To standardise with both the non-reactive and HPC DSTs, the 95th percentile was chosen, which was calculated to be 73 µg/cm^2. By applying Equation (1) and using the performance figures for the subset of the expanded dataset classified as non-HPC (sensitivity = 81%, specificity = 65%), the probability that an unknown chemical classed as reactive, but non-HPC would be a sensitiser was calculated to be 36.4%. The probability of a reactive (non-HPC) chemical having an experimental EC3 value less than 73 µg/cm^2 is therefore 0.05 × 0.364 = 0.0182, meaning that the updated non-reactive DST is 98.2% protective.

Finally, the 95th percentile was also taken from the gamma distribution of the 146 HPC sensitisers, giving an updated HPC DST of 1.0 µg/cm^2. This value is similar to, and only 1.5-fold smaller than, the previously published HPC DST of 1.5 µg/cm^2. By applying Equation (1) and using the performance figures for the subset of the expanded dataset classified as HPC (sensitivity = 99%, specificity = 37%), the probability that an unknown chemical classed as HPC would be a sensitiser can be

calculated to be 28.1%. Therefore, the probability of an HPC chemical having an experimental EC3 value less than 1.0 μg/cm^2 is 0.05 × 0.281 = 0.0141, meaning that the updated non-reactive DST is 98.6% protective.

Across the expanded DST dataset, there were 16 sensitisers that have an EC3 value below their corresponding DST, which were therefore investigated further to see if the discrepancies could be rationalised (Table 3). Of the three non-reactive sensitisers with an EC3 value below the updated non-reactive DST, two are suspected of containing strongly sensitising impurities, while one represents a potential knowledge gap in Derek Nexus.

For the 10 reactive (non-HPC) sensitisers that have an EC3 value below the reactive (non-HPC) DST, a rationale for the lack of HPC prediction by Derek Nexus could be provided for four of them: one is likely to be a false positive response in the LLNA, one has a potency value which is likely to be over-estimated based on inappropriate extrapolation from the dose-response data (Ryan et al., 2007), one has an unreliable EC3 value based on a comparison to known analogues, and one represents complex chemistries that are not immediately obvious to a human expert or an *in silico* expert system alike. The remaining six chemicals represent cases where a human expert would be likely to conclude HPC, but Derek Nexus has not; five of these were based on the use of HPC rule 9 due to structural complexity. The inability to easily encode such a rule into an *in silico* expert system is one of the limitations of such automation, as described previously. This could be mitigated by including an additional human expert review of those chemicals classified as reactive but not HPC, to judge whether they should be reclassified as HPC based on any inherent structural complexity.

Three HPC sensitisers have such an extreme skin sensitisation potency that their EC3 value is below the updated HPC DST. Upon investigation, two of these have a potency value that may be over-estimated or miscalculated based on extrapolating from the dose-response data (Ryan et al., 2007), supported by additional potency estimates using the probit extrapolation method (Roberts, 2015). The remaining chemical, 4-ethoxymethylene-2-phenyl-2-oxazolin-5-one, has a remarkably low EC3 value of 0.35 μg/cm^2 derived by linear interpolation. It has been suggested that the extreme potency of this chemical is due to its particular reaction chemistry, which is both very fast and diverse (Natsch et al., 2010).

A fuller understanding of the individual reasons why these murine sensitisers are mispredicted may emerge over time as new chemistry knowledge and/or experimental data becomes available, which in turn should enable further refinement of the DST values.

Finally, there are two questions that a user of the DST approach may well ask of this research:

1.) Which set of DST values should be used, given that the updated values differ to the previously published values?
2.) How should a chemical be assigned to its correct reactivity class, to select the appropriate DST?

In answer to the first question, the similarity of the updated DST values to the original values (between 1.1- and 1.5-fold difference) shows that the broad trends previously published, that enabled the likely potency of non-reactive, reactive, and HPC chemicals to be differentiated, have not changed significantly. Based on the available LLNA data, the potency range of known sensitisers can encompass at least five orders of magnitude (Gerberick et al., 2005; Kern et al., 2010). Practically, the small changes to the DST values based on the expanded dataset can be considered physiologically insignificant in a Quantitative Risk Assessment. While the non-reactive and HPC DST values have reduced slightly in this updated analysis, meaning that their use will result in slightly stricter thresholds within a risk assessment, the recommendation would be to employ these DST values going forwards. This is based on the consideration that, as the dataset they are derived from is three times larger than the original DST dataset (and contains twice as many sensitisers), the thresholds derived from it are likely to be more representative of the wider chemical universe, and therefore more robust and ultimately more protective of human health.

In answer to the second question, this study has described how the reactivity classification can be automated using the skin sensitisation and HPC alerts encoded in the *in silico* expert system Derek Nexus. It is hoped that this will be a useful option for many who want to use the DST approach, although it is of course not the only option. A human expert can equally apply the skin sensitisation reactivity domains and HPC rules to a new chemical, other *in silico* tools are available to classify sensitisation reactivity, and if necessary experimental chemistry can be applied to assess the nature and degree of reactivity for a difficult-to-classify substance. In practice, *in silico* expert knowledge and human expert judgment often come to the same conclusion (as demonstrated in this study). However, rather than seeing these different approaches to reactivity classification as mutually exclusive, they could in fact be complementary; with *in silico* expert knowledge, human expert judgement and experimental chemistry all being viable options to be used alone, or in combination, to help assign the most appropriate DST value to a chemical of unknown reactivity and unknown sensitisation potential.

4. Conclusions

This study has shown that Derek Nexus is a suitable *in silico* reactivity classifier which can be used to calculate Dermal Sensitisation Thresholds. The performance of Derek Nexus against the original DST dataset was very similar to that of human experts, and the recalculated DSTs based on this dataset were also similar, once a particularly impactful false positive outlier (hexyl salicylate) had been removed. As implemented in Derek Nexus, the updated HPC alerts are shown to be suitable for identifying high potency sensitisers.

The expanded DST dataset created during this study showed a similar overall EC3 distribution to that of the original DST dataset, despite containing twice as many sensitisers. This similarity demonstrates that the reported gamma distribution continues to be representative of the potency of known skin sensitisers found in the public domain.

By using Derek Nexus to classify the sensitisers in the expanded DST dataset, three updated DSTs have been derived: a non-reactive DST of 710 μg/cm^2 based on 79 chemicals, a reactive (non-HPC) DST of 73 μg/cm^2 based on 331 chemicals, and an HPC DST of 1.0 μg/cm^2 based on 146 chemicals. The proximity of the updated non-reactive and HPC DSTs to the previously published values highlights the robustness of these thresholds, even though the expanded DST dataset contains over three times as many chemicals. The updated reactive (non-HPC) DST complements the previously published reactive DST, as although they are derived in different manners, the actual values are very similar to each other and should be able to be used interchangeably. In those rare cases where a chemical in the expanded dataset has an EC3 value lower than the corresponding DST, this can often be rationalised (e.g., due to impurities in the sample tested or unreliable interpretation of the LLNA dose-response data), although occasional gaps within the knowledge base of the *in silico* expert system were also observed. Calculations showed that the probability of a chemical being less potent than the relevant DST was very high in all cases (98.2–99.7%), thus giving confidence that this updated approach remains highly protective of human health.

The analysis of the expanded dataset supports the robustness of the DST thresholds and should increase confidence in the protectiveness of these methods. Furthermore, the use of Derek Nexus as an *in silico* reactivity classifier provides a reproducible, scalable and accessible approach to applying the updated DSTs within a Quantitative Risk Assessment for skin sensitisation.

Funding body information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Martyn L. Chilton: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Writing – original draft, Writing – review & editing. **Anne Marie Api:** Data curation, Methodology, Writing – review & editing. **Robert S. Foster:** Data curation, Software. **G. Frank Gerberick:** Data curation, Methodology, Writing – review & editing. **Maura Lavelle:** Data curation, Methodology, Writing – review & editing. **Donna S. Macmillan:** Conceptualization, Data curation, Methodology, Software, Supervision, Writing – review & editing. **Mihwa Na:** Data curation, Methodology, Writing – review & editing. **Devin O'Brien:** Data curation, Methodology. **Catherine O'Leary-Steele:** Data curation, Software. **Mukesh Patel:** Data curation, Software. **David J. Ponting:** Data curation, Software. **David W. Roberts:** Data curation, Formal analysis, Methodology, Writing – review & editing. **Robert J. Safford:** Data curation, Methodology, Writing – review & editing. **Rachael E. Tennant:** Data curation, Software.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Martyn Chilton reports a relationship with Lhasa Limited that includes: employment. Anne Marie Api reports a relationship with Research Institute for Fragrance Materials that includes: employment. Robert Foster reports a relationship with Lhasa Limited that includes: employment. Frank Gerberick reports a relationship with Research Institute for Fragrance Materials that includes: consulting or advisory. Maura Lavelle reports a relationship with Research Institute for Fragrance Materials that includes: employment. Donna Macmillan reports a relationship with Lhasa Limited that includes: employment. Mihwa Na reports a relationship with Research Institute for Fragrance Materials that includes: employment. Devin O'Brien reports a relationship with Research Institute for Fragrance Materials that includes: employment. Catherine O'Leary-Steele reports a relationship with Lhasa Limited that includes: employment. Mukesh Patel reports a relationship with Lhasa Limited that includes: employment. David Ponting reports a relationship with Lhasa Limited that includes: employment. David Roberts reports a relationship with Research Institute for Fragrance Materials that includes: consulting or advisory. Robert Safford reports a relationship with Research Institute for Fragrance Materials that includes: consulting or advisory. Rachael Tennant reports a relationship with Lhasa Limited that includes: employment.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.yrtph.2022.105200.

References

Api, A.M., Basketter, D., Bridges, J., Cadby, P., Ellis, G., Gilmour, N., Greim, H., Griem, P., Kern, P., Khaiat, A., O'Brien, J., Rustemeyer, T., Ryan, C., Safford, B., Smith, B., Vey, M., White, I.R., 2020. Updating exposure assessment for skin sensitization quantitative risk assessment for fragrance materials. Regul. Toxicol. Pharmacol. 118, 104805 https://doi.org/10.1016/j.yrtph.2020.104805.

Api, A.M., Basketter, D., Lalko, J., 2015. Correlation between experimental human and murine skin sensitization induction thresholds. Cutan. Ocul. Toxicol. 34, 298–302. https://doi.org/10.3109/15569527.2014.979425.

Api, A.M., Basketter, D.A., Cadby, P.A., Cano, M.-F., Ellis, G., Gerberick, G.F., Griem, P., McNamee, P.M., Ryan, C.A., Safford, R., 2008. Dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. Regul. Toxicol. Pharmacol. 52, 3–23. https://doi.org/10.1016/j.yrtph.2007.10.008.

Aptula, A.O., Patlewicz, G., Roberts, D.W., 2005. Skin sensitization: reaction mechanistic applicability domains for Structure−Activity relationships. Chem. Res. Toxicol. 18, 1420–1426. https://doi.org/10.1021/tx050075m.

Ashikaga, T., Sakaguchi, H., Sono, S., Kosaka, N., Ishikawa, M., Nukada, Y., Miyazawa, M., Ito, Y., Nishiyama, N., Itagaki, H., 2010. A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). Altern. Lab. Anim. 38, 275–284. https://doi.org/10.1177/026119291003800403.

Basketter, D.A., Alepee, N., Ashikaga, T., Barroso, J., Gilmour, N., Goebel, C., Hibatallah, J., Hoffmann, S., Kern, P., Martinozzi-Teissier, S., Maxwell, G., Reisinger, K., Sakaguchi, H., Schepky, A., Tailhardat, M., Templier, M., 2014. Categorization of chemicals according to their relative human skin sensitizing potency. Dermatitis 25, 11–21. https://doi.org/10.1097/DER.0000000000000003.

Basketter, D.A., Clapp, C., Jefferies, D., Safford, B., Ryan, C.A., Gerberick, F., Dearman, R.J., Kimber, I., 2005. Predictive identification of human skin sensitization thresholds. Contact Dermatitis 53, 260–267. https://doi.org/10.1111/J.0105-1873.2005.00707.X.

Bergström, M.A., Andersson, S.I., Broo, K., Luthman, K., Karlberg, A.-T., 2008. Oximes: metabolic activation and Structure−Allergenic activity relationships. J. Med. Chem. 51, 2541–2550. https://doi.org/10.1021/jm701092n.

Berthold, M.R., Cebron, N., Dill, F., Gabriel, T.R., Kötter, T., Meinl, T., Ohl, P., Thiel, K., Wiswedel, B., 2009. Knime - the Konstanz information miner. In: Preisach, C., Burkhardt, H., Schmidt-Thieme, L., Decker, R. (Eds.), ACM SIGKDD Explorations Newsletter. Springer, Berlin, Heidelberg, p. 26. https://doi.org/10.1145/1656274.1656280.

Cramer, G.M., Ford, R.A., Hall, R.L., 1976. Estimation of toxic hazard—a decision tree approach. Food Chem. Toxicol. 16, 255–276. https://doi.org/10.1016/S0015-6264(76)80522-6.

Gerberick, F.G., Ryan, C.A., Kern, P.S., Schlatter, H., Dearman, R.J., Kimber, I., Patlewicz, G.Y., Basketter, D.A., 2005. Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. Dermatitis 16, 157–202. https://doi.org/10.1097/01206501-200512000-00002.

Heller, S.R., McNaught, A., Pletnev, I., Stein, S., Tchekhovskoi, D., 2015. InChI, the IUPAC international chemical identifier. J. Cheminf. 7, 23. https://doi.org/10.1186/s13321-015-0068-4.

Kern, P.S., Gerberick, F.G., Ryan, C.A., Kimber, I., Aptula, A., Basketter, D.A., 2010. Local lymph node data for the evaluation of skin sensitization alternatives: a second compilation. Dermatitis 21, 8–32. https://doi.org/10.2310/6620.2009.09038.

Kimber, I., Gerberick, G.F., Basketter, D.A., 2017. Quantitative risk assessment for skin sensitization: success or failure? Regul. Toxicol. Pharmacol. 83, 104–108. https://doi.org/10.1016/j.yrtph.2016.11.020.

Kroes, R., Renwick, A.G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., Schilter, B., Schlatter, J., van Schothorst, F., Vos, J.G., Würtzen, G., 2004. Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. Food Chem. Toxicol. 42, 65–83. https://doi.org/10.1016/j.fct.2003.08.006.

Lhasa Limited, 2021. Derek Nexus v6.1.1 [WWW Document]. URL. https://www.lhasalimited.org/products/derek-nexus.htm, 11.22.19.

Loveless, S.E., Ladies, G.S., Gerberick, G.F., Ryan, C.A., Basketter, D.A., Scholes, E.W., House, R.V., Hilton, J., Dearman, R.J., Kimber, I., 1996. Further evaluation of the local lymph node assay in the final phase of an international collaborative trial. Toxicology 108, 141–152. https://doi.org/10.1016/0300-483X(95)03279-O.

Miralles, J., Escudero, A., Carbonell, A., Martínez, A., Fernández, E., Cardona, P., 2015. Allergic contact dermatitis from ethylhexyl salicylate. J Investig. Allergol. Clin. Immunol. 25, 55–82.

Montelius, J., Wahlkvist, H., Boman, A., Wahlberg, J.E., 1998. Murine local lymph node assay for predictive testing of allergenicity: two irritants caused significant proliferation. Acta Derm. Venereol. 78, 433–437. https://doi.org/10.1080/000155598442728.

Mortz, C.G., Thormann, H., Goossens, A., Andersen, K.E., 2010. Allergic contact dermatitis from ethylhexyl salicylate and other salicylates. Dermatitis 21, E7–E10. https://doi.org/10.2310/6620.2010.09090.

Munro, I.C., Ford, R.A., Kennepohl, E., Sprenger, J.G., 1996. Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. Food Chem. Toxicol. 34, 829–867. https://doi.org/10.1016/S0278-6915(96)00049-X.

Natsch, A., Gfeller, H., Kuhn, F., Granier, T., Roberts, D.W., 2010. Chemical basis for the extreme skin sensitization potency of (E)-4-(Ethoxymethylene)-2-phenyloxazol-5(4H)-one. Chem. Res. Toxicol. 23, 1913–1920. https://doi.org/10.1021/tx1002707.

NICEATM, 2010. NICEATM Murine Local Lymph Node Assay (LLNA) Database [WWW Document]. URL. https://ntp.niehs.nih.gov/whatwestudy/niceatm/test-method-evaluations/immunotoxicity/nonanimal/index.html#NICEATM-Murine-Local-Lymph-Node-Assay-LLNA-Database, 1.3.20.

Nishijo, T., Api, A.M., Gerberick, G.F., Miyazawa, M., Roberts, D.W., Safford, R.J., Sakaguchi, H., 2020. Application of the dermal sensitization threshold concept to chemicals classified as high potency category for skin sensitization assessment of ingredients for consumer products. Regul. Toxicol. Pharmacol. 117, 104732 https://doi.org/10.1016/j.yrtph.2020.104732.

Roberts, D.W., 2015. Estimating skin sensitization potency from a single dose LLNA. Regul. Toxicol. Pharmacol. 71, 437–443. https://doi.org/10.1016/j.yrtph.2015.01.009.

Roberts, D.W., Api, A.M., 2018. Chemical applicability domain of the local lymph node assay (LLNA) for skin sensitisation potency. Part 4. Quantitative correlation of LLNA

potency with human potency. Regul. Toxicol. Pharmacol. 96, 76–84. https://doi.org/10.1016/j.yrtph.2018.04.022.

Roberts, D.W., Api, A.M., Safford, R.J., Lalko, J.F., 2015. Principles for identification of high potency category chemicals for which the dermal sensitisation threshold (DST) approach should not be applied. Regul. Toxicol. Pharmacol. 72, 683–693. https://doi.org/10.1016/j.yrtph.2015.03.001.

Roberts, D.W., Aptula, A., Api, A.M., 2017. Structure–potency relationships for epoxides in allergic contact dermatitis. Chem. Res. Toxicol. 30, 524–531. https://doi.org/10.1021/acs.chemrestox.6b00241.

Ryan, C.A., Chaney, J.G., Gerberick, G.F., Kern, P.S., Dearman, R.J., Kimber, I., Basketter, D.A., 2007. Extrapolating local lymph node assay EC3 values to estimate relative sensitizing potency. Cutan. Ocul. Toxicol. 26, 135–145. https://doi.org/10.1080/15569520701212258.

Safford, R.J., 2008. The Dermal Sensitisation Threshold- a TTC approach for allergic contact dermatitis. Regul. Toxicol. Pharmacol. 51, 195–200. https://doi.org/10.1016/j.yrtph.2008.02.010.

Safford, R.J., Api, A.M., Roberts, D.W., Lalko, J.F., 2015. Extension of the Dermal Sensitisation Threshold (DST) approach to incorporate chemicals classified as reactive. Regul. Toxicol. Pharmacol. 72, 694–701. https://doi.org/10.1016/j.yrtph.2015.04.020.

Safford, R.J., Aptula, A.O., Gilmour, N., 2011. Refinement of the Dermal Sensitisation Threshold (DST) approach using a larger dataset and incorporating mechanistic chemistry domains. Regul. Toxicol. Pharmacol. 60, 218–224. https://doi.org/10.1016/j.yrtph.2011.03.009.

Shaw, D.W., 2006. Allergic contact dermatitis from octisalate and cis-3-hexenyl salicylate. Dermatitis 17, 152–155. https://doi.org/10.2310/6620.2006.05046.

Singh, M., Beck, M.H., 2007. Octyl salicylate: a new contact sensitivity. Contact Dermatitis 56. https://doi.org/10.1111/J.1600-0536.2007.00942.X, 48–48.

Slatter, D.A., Bolton, C.H., Bailey, A.J., 2000. The importance of lipid-derived malondialdehyde in diabetes mellitus. Diabetologia 43, 550–557. https://doi.org/10.1007/s001250051342.

Spielmann, H., Hoffmann, S., Liebsch, M., Botham, P., Fentem, J.H., Eskes, C., Roguet, R., Cotovio, J., Cole, T., Worth, A., Heylings, J., Jones, P., Robles, C., Kandárová, H., Gamer, A., Remmele, M., Curren, R., Raabe, H., Cockshott, A., Gerner, I., Zuang, V., 2007. The ECVAM international validation study on in vitro tests for acute skin irritation: report on the validity of the EPISKIN and EpiDerm assays and on the skin integrity function test. Altern. Lab. Anim. 35, 559–601. https://doi.org/10.1177/026119290703500614.

Wei, Q., Dunbrack, R.L., 2013. The role of balanced training and testing data sets for binary classifiers in bioinformatics. PLoS One 8, e67863. https://doi.org/10.1371/journal.pone.0067863.