scientific reports

Check for updates

OPEN

Quantitative assessment of sensitizing potency using a dose–response adaptation of GARDskin

Robin Gradin, Andy Forreryd, Ulrika Mattson, Anders Jerre & Henrik Johansson✉

Hundreds of chemicals have been identified as skin sensitizers. These are chemicals that possess the ability to induce hypersensitivity reactions in humans, giving rise to a condition termed allergic contact dermatitis. The capacity to limit hazardous exposure to such chemicals depends upon the ability to accurately identify and characterize their skin sensitizing potency. This has traditionally been accomplished using animal models, but their widespread use offers challenges from both an ethical and a scientific perspective. Comprehensive efforts have been made by the scientific community to develop new approach methodologies (NAMs) capable of replacing in vivo assays, which have successfully yielded several methods that can identify skin sensitizers. However, there is still a lack of new approaches that can effectively measure skin sensitizing potency. We present a novel methodology for quantitative assessment of skin sensitizing potency, which is founded on the already established protocols of the GARDskin assay. This approach analyses dose–response relationships in the GARDskin assay to identify chemical-specific concentrations that are sufficient to induce a positive response in the assay. We here compare results for 22 skin sensitizers analyzed using this method with both human and LLNA potency reference data and show that the results correlate strongly and significantly with both metrics ($r_{LLNA} = 0.81$, $p = 9.1 \times 10^{-5}$; $r_{Human} = 0.74$, $p = 1.5 \times 10^{-3}$). In conclusion, the results suggest that the proposed GARDskin dose–response methodology provides a novel non-animal approach for quantitative potency assessment, which could represent an important step towards reducing the need for in vivo experiments.

Contact allergens are chemicals that possess the intrinsic potential to induce skin sensitization resulting in allergic contact dermatitis (ACD). The mechanisms of skin sensitization have been extensively reviewed previously[1–4] and are summarized in an Adverse Outcome Pathway (AOP)[5]. Briefly, they include an asymptomatic induction phase involving priming of an adaptive immune response and clonal expansion of chemical specific T-cells resulting in allergic sensitization. If the sensitized subject is exposed subsequently to the same chemical, a cutaneous inflammatory reaction will be provoked that is described clinically as ACD.

Within the regulatory context, important tools for assessment of skin sensitization potential include the Guinea pig (GP) assays described in OECD TG 406[6] [(GP Maximization Test (GPMT)[7] and the Buehler Occluded Patch Test[8,9]], and the murine Local Lymph Node Assay (LLNA, TG 429)[10]. The LLNA is often the preferred method and provides several advantages compared with GP assays in terms of improved animal welfare, increased predictivity, and ability to provide a continuous estimate of sensitizing potency, as the chemical concentration required to elicit a threefold proliferation of T-lymphocytes (EC3 value) in the lymph node draining the site of topical application[11,12] can be identified by interpolation from the dose–response curve. Moreover, a suite of mechanistically based new approach methodologies (NAMs), which are non-animal-based assays, have recently become available for hazard identification of skin sensitizers. Their development has been motivated by legislative mandates, such as the revision of Annex VII of the REACH regulation[13], which makes non-animal testing the default requirement compared to animal tests within certain sectors. At the time of writing, seven such methods have been formally validated and incorporated into globally accepted test guidelines by OECD (OECD TG 442C[14], D[15] and E[16]).

A challenge for the complete replacement of animal studies for skin sensitization assessment, however, is that the current methods adopted by the OECD have thus far been validated only for hazard identification, and not

SenzaGen AB, Medicon Village, Scheelevägen 2, 22381 Lund, Sweden. ✉email: henrik.johansson@senzagen.com

for assessment of sensitizing potency, which, as indicated above, is a key requirement for effective risk management. Significant efforts have been made to develop in vitro and in silico methods as well as defined approaches (DA) for potency assessments, of which a majority focus on skin sensitizer GHS sub-categorization. However, as evident from a recent review article, this has proven to be a significant challenge, and the current state-of-the art strategies for such classifications exhibit predictive performances ranging between 55 and 69%[17]. Recently, a modified version of the Direct Peptide Reactivity Assay (DPRA, OECD TG 442 C[14], the kinetic DPRA (kDPRA)[18], has been described in a series of publications and has initially shown promising results for GHS potency sub-categorization[19–21], but it still remains to be determined how well this approach performs on external data with an a priori defined classification threshold. Furthermore, to this end, development and regulatory implementation of NAM based strategies capable of providing continuous potency data beyond the discrete GHS subcategories also remain a topic of high priority.

GARD (OECD TGP 4.106) is an in vitro testing platform that brings novel elements to the field of regulatory toxicology by monitoring transcriptional patterns of biomarker signatures in a human dendritic-like cell line and provides machine-learning assisted classifications. The GARDskin assay monitors genes in a biomarker signature involved in immunologically associated pathways relevant to several key events (KE) in the AOP to arrive at mechanistically-based classifications[22]. The GARDpotency assay monitors a complementary biomarker signature[23] and has been applied successfully for sub-categorization of sensitizers according to the GHS system. The assays were recently evaluated in a formal validation study, and reported cumulative predictive performances of 94% and 88%, for hazard identification[24] and GHS potency sub-categorization[25], respectively.

The present paper describes a novel adaption of the GARDskin protocol to support dose–response analyses to provide a quantitative and continuous measurement of skin sensitizing potency. The proposed strategy determines the minimal concentration required to exceed the binary classification threshold in GARDskin ($DV \geq 0$), here referred to as the cDV_0 concentration. Experimentally derived cDV_0 values from the testing of a total of 29 chemicals are presented, together with their associated dose–response curves. Comparisons with relevant potency metrics reveal statistically significant correlations with both LLNA EC3 and human No Observed Effect Levels (NOEL).

Results

GARDskin dose–response: methodological description. The conventional GARDskin protocol assays chemicals at single concentrations and performs hazard classifications into the categories: *skin sensitizer* and *non-sensitizer*. Although results from the assay are binary, it has been noted that the observed effects seem to be dose-dependent, demonstrating an inverse correlation between concentrations at which chemicals have been assayed, and their relative sensitizing potency. However, by itself, the GARD input concentration may not be sufficiently predictive, as observed previously[25]. Therefore, to further explore the relationship between chemical exposure concentrations and GARDskin responses, and its subsequent association with sensitizing potency, dose–response measurements were incorporated into the GARDskin protocol. Specifically, it was hypothesized that the lowest concentration required to exceed the binary classification threshold ($DV \geq 0$) and generate a positive classification, referred to as the cDV_0 concentration, would be informative of sensitizing potency. As such, the proposed dose–response measurements would also be aligned with common toxicological principles for potency assessments, where the response value (DV), the binary threshold ($DV \geq 0$), and the derived cDV_0 concentrations may be viewed as an analogue to the response value (stimulation index (SI)), the binary threshold ($SI \geq 3$), and the EC3 concentrations in the LLNA assay, respectively. To evaluate this premise, a total of 29 chemicals, comprised of 7 non-sensitizers and 22 skin sensitizers of varying potency, were tested in a titrated range of concentrations. Given that the experimental setup included several novel elements, prior knowledge of the dose–response relationships' characteristics were lacking. Therefore, the exposure concentrations were selected to evenly sample several concentrations of each chemical instead of favoring a few doses with replicates. This approach was preferred due to the reduced risk of selecting poor concentrations which could have resulted in poor estimates of the dose–response relationships[26,27]. The standard GARDskin pipeline was used to generate response values, by applying the SVM model to assign DVs to each chemical and concentration, based on the expression levels of the genes in the GARDskin biomarker signature. For each tested chemical, the assayed concentrations were plotted against the responses, i.e., the DVs, as illustrated in individual dose–response plots in Fig. 1.

Dose–response relationships and cDV_0 estimation. Following inspection of the acquired data, log-logistic models were fitted to each chemical to attempt to model the dose–response relationships. No-effect tests were applied to assess the presence of concentration-dependent increases in responses. The p-value for respective chemical and test is described in Table 1. As can be seen, the p-values for salicylic acid, xylene, 1-butanol, glycerol, octanoic acid, vanillin, trans-anethol, and kanamycin sulfate, agree with observations of the raw data (Fig. 1) and suggest that these chemicals did not induce increased responses in the assay over the considered concentrations. The results are further suggestive of weak responses for benzyl salicylate and benzocaine. Examination of the data showed that benzocaine did not exhibit any positive decision values, though an upward trend at the highest concentrations could potentially still be recognized, which would explain the relatively low p-value associated with the model's fit. The two highest concentrations of benzyl salicylate generated positive responses, though with relatively low magnitudes. Remaining chemicals exhibited more distinct dose–response relationships with more significant p-values from the no-effect tests, suggesting quantifiable effects on the response values.

Next, cDV_0 values were estimated for each chemical. Two different approaches were considered for the calculations. First, the concentrations were estimated using the already fitted log-logistic models by determining the

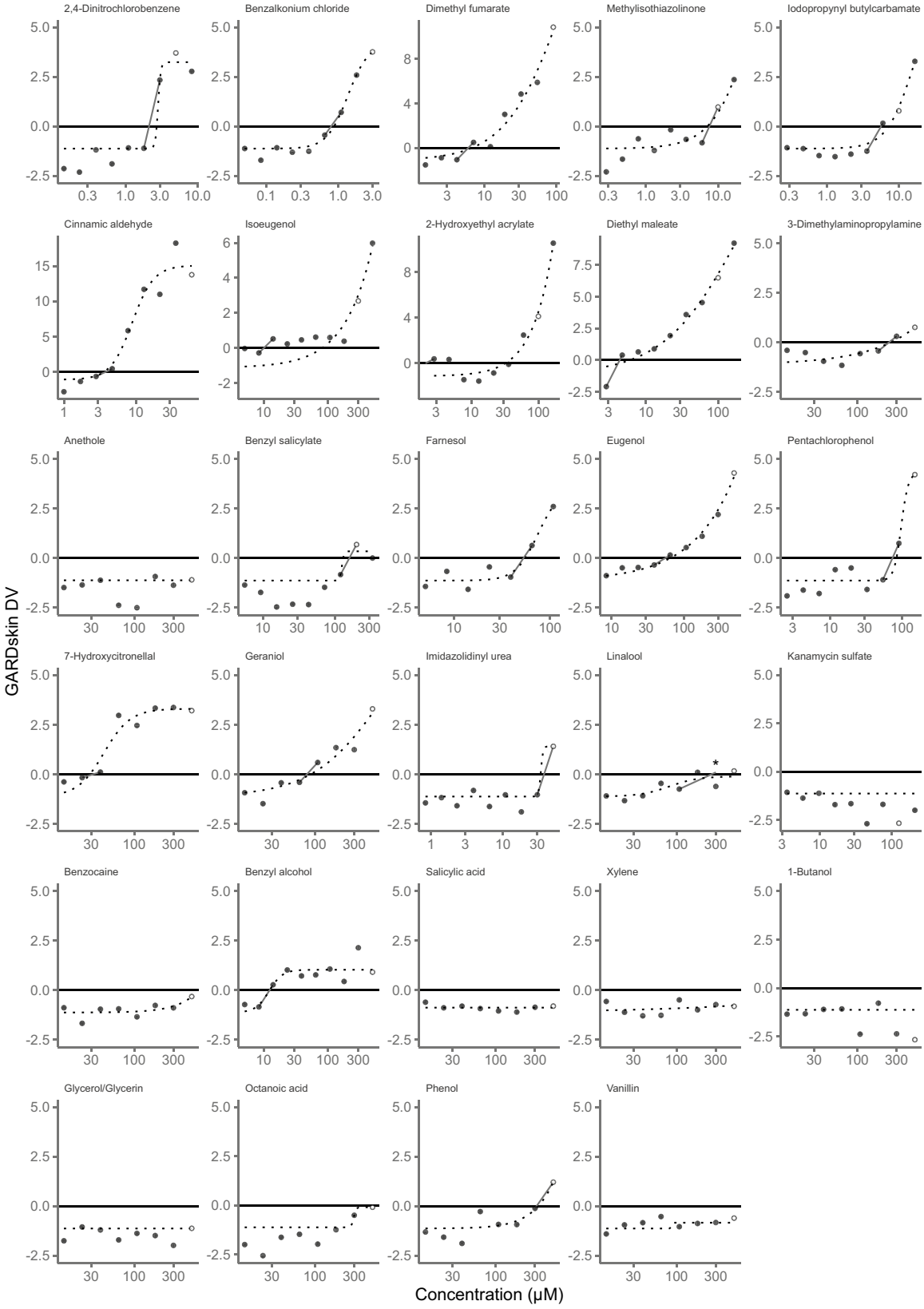

Figure 1. Dose–response relationships for the examined chemicals. Points represent measured response values (decision values) at respective concentrations. Points with white centers represent the GARD input concentrations. Dotted lines describe the log-logistic models, and the solid lines describe the linear interpolations. *The running median was used as positive interpolation point to reduce the potential impact of noise.

No	Chemical	CAS	LLNA EC3 (%)[a]	NOEL (μg/cm²)[b]	GARD input c. (μM)	cDV_0 (μM)	cDV_0 Log-Logistic (μM) (95% CI)	cDV_0 (mg/L)	P No-effect
1	2,4-Dinitrochlorobenzene	97-00-7	0.06	8.8	5	2.19	2.71 (2.00, 3.77)	0.443	1.05×10^{-5}
2	Benzalkonium chloride	8001-54-5	0.1	ND	3	0.826	0.87 (0.70, 1.05)	0.350	8.15×10^{-9}
3	Dimethyl fumarate	624-49-7	0.35	88	90	6.06	6.58 (2.88, 12.1)	0.974	1.71×10^{-7}
4	Methylisothiazolinone	2682-20-4	0.4[28]	15	10	7.85	7.12 (2.23, 13.5)	0.904	5.32×10^{-4}
5	Iodopropynyl butylcarbamate	55406-53-6	0.9	ND	10	5.74	7.24 (5.34, 9.78)	1.61	7.16×10^{-7}
6	Cinnamic aldehyde	104-55-2	1.15	591	60	3.97	3.64 (2.05, 6.64)	0.524	2.21×10^{-6}
7	Isoeugenol	97-54-1	1.35	69	300	10.4	81.8 (NA, NA)	1.70	8.62×10^{-4}
8	2-Hydroxyethyl acrylate	818-61-1	1.56	ND	100	NA	NA	NM	1.50×10^{-6}
9	Diethyl maleate	141-05-9	2.1[29]	1600	100	4.38	5.97 (2.20, 12.1)	0.753	2.76×10^{-7}
10	3-Dimethylaminopropylamine	109-55-7	2.2	ND	500	251	257 (107, NA)	25.7	6.72×10^{-3}
11	trans-Anethole	4180-23-8	2.7[30]	5510[31]	500	NS	NS	NS	1.00
12	Benzyl salicylate	118-58-1	2.85	17,717	200	164	130 (97.9, NA)	37.4	7.74×10^{-2}
13	Farnesol	4602-84-0	4.8	2755	500	54.4	55.0 (26.4, 92.9)	12.1	1.09×10^{-4}
14	Eugenol	97-53-0	12.9	1938	500	56.6	64.6 (57.8, 72.6)	9.29	1.96×10^{-9}
15	Pentachlorophenol	87-86-5	20	2155	150	75.4	83.3 (52.8, 137)	20.1	7.11×10^{-6}
16	7-Hydroxycitronellal	107-75-5	22.2	2953	500	33.1	29.0 (14.3, 43.1)	5.70	4.85×10^{-5}
17	Geraniol	106-24-1	23.2	3875	500	82.6	88.9 (39.8, 183)	12.7	3.68×10^{-5}
18	Imidazolidinyl urea	39236-46-9	24	2000	50	38.5	33.4 (25.6, 46.0)	14.9	3.22×10^{-4}
19	Linalool	78-70-6	30.4	13,793	500	279	NS (71.6, NA)	43.0	9.54×10^{-3}
20	Kanamycin sulfate	70560-51-9	NS	1874	125	NS	NS	NS	1.00
21	Benzocaine	94-09-7	NS	2000	500	NS	NS (406, NA)	NS	5.76×10^{-2}
22	Benzyl alcohol	100-51-6	NS	5906[31]	500	12.6	12.4 (6.32, 22.8)	1.37	1.60×10^{-3}
23	Salicylic acid	69-72-7	12.2	NS	500	NS	NS	NS	1.00
24	Xylene	1330-20-7	95.8	NS	500	NS	NS	NS	0.810
25	1-Butanol	71-36-3	NS	NS	500	NS	NS	NS	1.00
26	Glycerol	56-81-5	NS	NS	500	NS	NS (181, NA)	NS	1.00
27	Octanoic acid	124-07-2	NS	NS	500	NS	NS (224, NA)	NS	0.555
28	Phenol	108-95-2	NS	NS	500	317	328 (150, NA)	29.9	2.44×10^{-3}
29	Vanillin	121-33-5	NS	1181	500	NS	NS	NS	1.00

Table 1. Chemicals used to assess the GARDskin dose–response methodology and their measures of skin sensitizing potency and estimated cDV_0 values. *NA* Value could not be defined, *ND* Data insufficient for defining a NOEL, *NS* Non-sensitizer. [a]Reference-data for LLNA EC3 values were, unless otherwise stated, obtained from Hoffman et al.[32]. [b]Reference-data for human NOEL values were, unless otherwise stated, obtained from Basketter et al.[33].

curves' x-axis intercepts. The log-logistic models also allowed for characterization of the uncertainty of respective cDV_0 estimate, by calculation of approximate 95% confidence intervals, see Table 1. The log-logistic models do, however, make certain assumptions regarding the underlying data and the relationship between concentrations and response values, which might not necessarily be valid, while simultaneously requiring a relatively complex fitting procedure. Therefore, linear interpolation between two points on adjacent sides of the decision border was considered as an alternative, simpler and less constrained, method for estimation of cDV_0. The derived concentration estimates from both approaches are compared in Fig. 2. As can be seen, both methods generally produced very similar concentrations. Three chemicals received inconsistent estimates when the two methodologies were compared. These were 2-hydroxyethyl acrylate, linalool, and isoeugenol. When examining the dose–response relationships for each of these chemicals, it can be seen that the observed discrepancy for 2-hydroxyethyl acrylate arose due to the two positive points present at the lowest exposure concentrations, giving an apparent U-shape to the overall response. Because of the unexpected appearance of these data, the chemical was excluded from further assessment. For linalool, a cDV_0 concentration was only definable with linear interpolation. For isoeugenol, the conflicting estimates followed from the shape of the dose–response curve, which seemed to deviate from the sigmoidal shape expected by the log-logistic models. Instead, the data suggested that the chemical was able to render positive decision values at moderate magnitudes already at low concentrations, which was followed by a plateau until concentrations reached > 180 μM, at which a second increase in the decision values was observed. Given this appearance, it can be argued that the locality of the linear interpolation provides a better estimate of the true cDV_0-value, since its fit is not constrained by a pre-determined shape of the curve. However, additional data would be useful for determining the induced dose–response curve's shape with more certainty. Because of

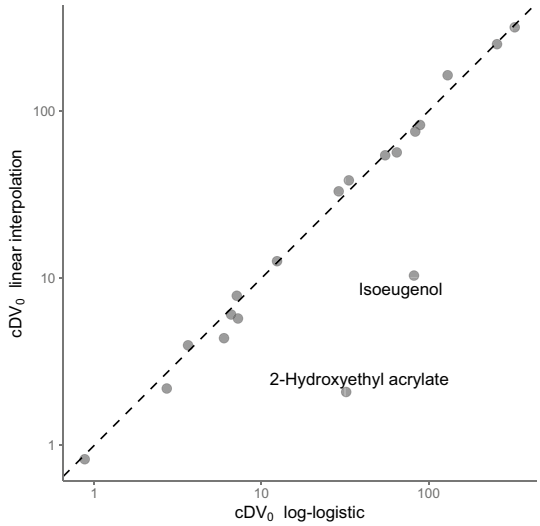

Figure 2. Comparison between the cDV_0 estimation procedures. The scatter plot describes cDV_0 values that were determined using either log-logistic modelling (x-axis) or linear interpolation (y-axis). The dashed line represents the identity line.

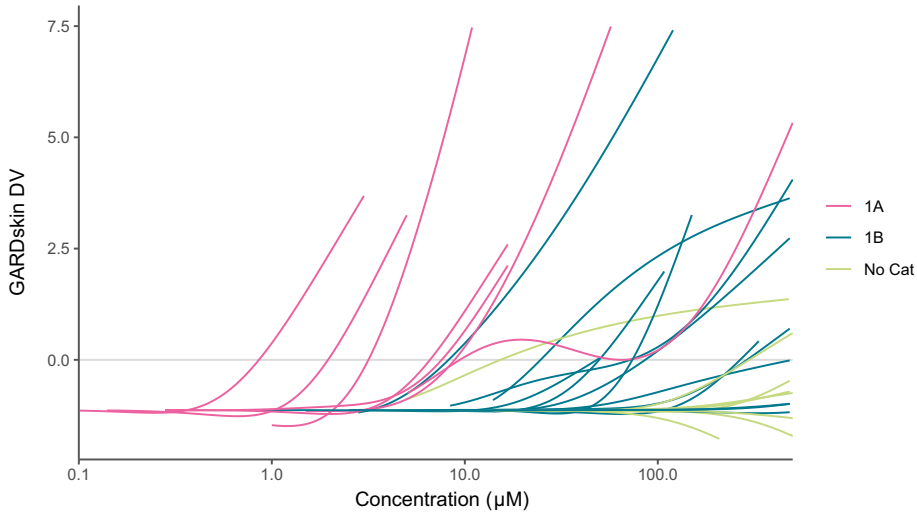

Figure 3. Visualization of acquired dose–response data. Each line represents a generalized additive model fitted to the dose–response data of a single chemical. The colors of the lines describe the GHS potency sub-categories, into which chemicals were assigned based on their LLNA results (1A: EC3 ≤ 2%; 1B: EC3 > 2%; No Cat: NS).

these observations, and the overall similarity between the results, the linear interpolation method was selected as the most appropriate technique for estimating cDV_0, and the results described below are based on these interpolated values, which are described in Table 1.

Comparison of cDV_0 with established potency measures. A visual representation of the experimentally derived dose–response curves for all chemicals are presented in Fig. 3, where the individual dose–response curves are colored according to their respective potency sub-category as defined by UN GHS (1A, 1B, no Cat) by the application of a 2% LLNA EC3 threshold to discriminate the stronger sensitizers from the weaker. As apparent by the figure, two of the LLNA non-sensitizers, benzyl alcohol and phenol, exhibited response-values sufficient to exceed the binary classification threshold (DV ≥ 0) at any of the assayed concentrations. The positive responses corresponded to several data points of benzyl alcohol, which is nevertheless generally considered a human skin sensitizer, and the highest evaluated concentration of phenol (500 μM). Further inspection of the figure and the generated dose–response curves clearly reveals a relationship between GHS potency and assayed concentrations, where stronger sensitizers (1A) required lower concentrations to exceed the binary classification threshold than did the weaker sensitizers (1B). Interestingly, with only a single exception for the chemical

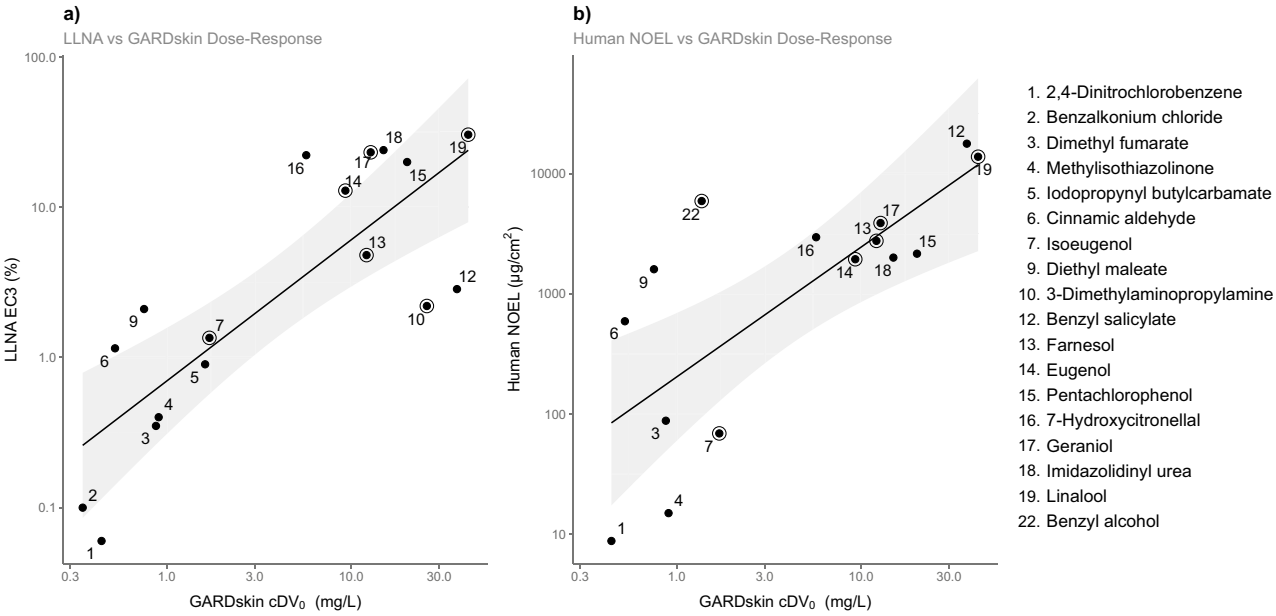

Figure 4. Performance assessment of GARDskin Dose–Response. Scatter plots displaying the relationship between estimated cDV_0 values and (**a**) LLNA EC3 values and (**b**) human NOEL values. The lines represent linear regression models fitted to the data, and the shaded areas describe the 95% confidence intervals of the fits. Encircled datapoints represent indirectly acting haptens. GARDskin cDV_0 is given in weight-based concentrations to facilitate comparison between potency measures.

diethyl maleate, which is a CLP 1A/1B borderline chemical, a complete resolution between the GHS potency classes was observed.

To further evaluate the hypothesis that cDV_0 values were informative of the chemicals' relative potencies, the correlations between cDV_0 and available potency measures were determined. First, the correlation with LLNA was considered. Figure 4 displays a scatter plot which suggests that chemicals ranked as strong sensitizers in the LLNA also seem to obtain lower cDV_0 values. The linear association between the two metrics was quantified using Pearson's correlation coefficient, which was determined to be 0.81 ($p = 9.1 \times 10^{-5}$; $n = 17$), which is indicative of a strong relationship, suggesting that GARDskin cDV_0 was significantly associated with sensitizing potency, as defined by LLNA EC3 values.

Second, the relationship between human NOEL values and cDV_0 values was considered and Fig. 4 shows the relationship between these two metrics. Similar to the LLNA-figure, a clear correlation can be observed. Indeed, quantification of the strength of the linear association resulted in a Pearson correlation coefficient of 0.74 ($p = 1.5 \times 10^{-3}$; $n = 15$), further supporting the observation that cDV_0 values were informative of sensitizing potency.

Finally, indirectly acting haptens, i.e., chemicals that require either abiotic or biotic activation prior to being able to induce sensitization reactions, are also displayed in Fig. 4. As can be seen, the data suggest that cDV_0 values are also informative of potency for these types of chemicals, suggesting that the GARDskin dose–response methodology is also able to assess indirectly acting haptens.

Discussion

The ability to identify and characterize chemical hazards is a requirement for effective risk assessment and risk management. Induction of skin sensitization is a threshold phenomenon[34], and the risk assessment procedure follow the general toxicological principles of quantitative risk assessments (QRA). The first step in this procedure is to derive a continuous prediction of skin sensitizing potency to define a concentration below which no reactions are expected to occur, a so-called No Expected Sensitization Induction Level (NESIL) value[35], for use as a point-of-departure (POD) in the QRA. While there have been considerable efforts within the scientific community toward the development of non-animal methods capable of providing such information, animal data, preferably from the LLNA, still remain an important source of information for quantitative potency assessments. For example, it can be used in regulatory contexts to derive discrete GHS sub-categorization[11] or in risk assessments where the NESIL value can be derived directly from the LLNA EC3 value, which correlates well with results from Human Repeated Insult Patch Testing (HRIPT)[36]. The ability to acquire such information from NAM technologies would nevertheless comprise an essential component in a non-animal toxicological toolbox for skin sensitizer assessment for the ultimate goal of replacing animal experimentation. In this context, it is worthwhile to remember that any assay developed towards a complex endpoint such as skin sensitization will likely be associated with errors, which also underscores the importance of effective post market surveillance programs for preservation of human health[37,38].

The GARDskin assay was developed for hazard identification of skin sensitizers and has been shown to be capable of accurately discriminating skin sensitizers from non-sensitizers[24]. In this study, we extend the original GARDskin method to include dose–response analysis for the purpose of extracting information appropriate for quantitative hazard characterization. Of the 22 potential skin sensitizers included in the analysis, 19 gave rise to positive response values at some of the assayed concentrations. The three chemicals that failed to induce any detectable responses were anethole, benzocaine, and kanamycin sulphate. It is worth noting that all of these chemicals are considered to be very weak skin sensitizers, and the expected classification labels for kanamycin and benzocaine are in fact ambiguous comparing human data to LLNA. The data recorded for 2-hydroxyethyl acrylate was ambiguous which hindered the chemical's further analysis. Additional data will be required to elucidate the dose–response curve's actual shape and characteristics. For each of the remaining sensitizers, the lowest concentration capable of rendering a positive prediction was estimated using linear interpolation, and the cDV_0 values were compared with potency measures available for respective chemical. It was found that cDV_0 correlated significantly with both LLNA EC3 ($r = 0.81$, $p = 9.1 \times 10^{-5}$) and human NOEL ($r = 0.74$, $p = 1.5 \times 10^{-3}$). The magnitudes of the correlation coefficients suggested that the associations between cDV_0 and both potency measures were strong, emphasizing the potential usefulness of the method. Moreover, the inclusion of several indirectly acting haptens among the chemicals indicated that GARDskin Dose–Response was also capable of generating meaningful cDV_0 values for these types of chemicals. The assessment of indirectly acting haptens are important as the ability of NAMs to assess such chemicals is a point of general concern[39].

It is also interesting to consider and compare the achieved performance figures with those obtained by other non-animal methods hitherto proposed for quantitative potency assessment. However, the number of overlapping chemicals among these assays are still minute, making such comparisons uncertain and not sufficiently informative to allow characterization of potential advantages or disadvantages of respective method. Therefore, we would like to postpone such comparisons until larger coherent sets of overlapping chemicals are available. Nevertheless, for the interested reader, performance measures for alternative methods on overlapping set of chemicals have been provided in Supplementary Data. Also, in the general context of assessing the results, it is necessary to acknowledge the difficulty in evaluating quantitative potency data due to imperfect reference data. Ideally, quantitative results from novel non-animal assays would be compared to well defined human-relevant potency references. However, such reference data are exceedingly rare and alternative measures have to be considered. In this study, we have included a comparison between cDV_0 and human NOEL. However, NOEL values are not a proper potency measure but corresponds to a concentration where no adverse effect has been observed, meaning that it is likely that concentrations exceeding NOEL might still not induce sensitization. And while it is true that they carry information of potency, they remain an imperfect source of reference.

There are several advantages to be had when an assay provides quantitative continuous potency results compared with qualitative or discrete ones, including the ability to rank chemicals based on their relative potencies', or the possibility to use the results as input in QRAs. This is a feature that separates the outcomes of the herein proposed GARDskin Dose–Response methodology and the previously described GARDpotency assay, which discriminates between strong and weak sensitizers based on the induced gene expression levels of a separate biomarker signature[23,25]. Thus, while the two assays might arguably provide some overlapping information, it would be expected that the GARDskin Dose–Response assay could provide relatively more information but at the cost of additional resources. Therefore, based on the information requirements of a particular test case, the two assays could still fill complementary data gaps. The generated cDV_0 values seem to correlate strongly with both examined potency measures, suggesting that the GARDskin Dose–Response assay could constitute an early but promising non-animal method for skin sensitizing hazard characterization. Assuming that the evidence described in this paper can be further reinforced on larger chemical sets, the proposed method could act as a direct alternative to the LLNA. The experimentally derived data would also be interpreted very similarly, which should make results accessible. However, it should again be noted that the examined dataset contains a limited number of chemicals, and additional data would indeed be necessary to increase the certainty of the observed association with sensitizing potency.

Another aspect of the analysis concerns the experimental protocol and the selection of assayed concentrations. Given that this was an initial study aimed towards examining the hypothesis that the lowest concentration capable of generating a positive GARDskin prediction would carry information of skin sensitizing potency, a few unknowns regarding the experimental procedures existed. For example, the characteristics of the responses and their dependence on the chemical exposure concentration was not known. Therefore, to increase the likelihood of observing response values in the vicinity of the decision boundary, it was prioritized to evaluate multiple concentrations in contrast to a few but replicated steps. This decision can also be justified by the non-controversial assumption that some smooth function describing the relationship between concentration and response exists (which is also reinforced by the relatively large number of assayed chemicals in GARDskin at utilized concentrations that show positive responses), which could potentially be more efficiently explored by an adequate number of sampled concentrations[26,27,40]. However, given these results, the selection of exposure concentrations could likely be optimized for future experiments to reduce uncertainties in cDV_0 estimates and to reduce time and cost expenditures. This analysis is, however, considered outside the scope of this publication since it could potentially require refinements as additional data are accumulated.

In conclusion, based on results presented in this paper, we argue that the proposed protocol, based on the validated protocol of GARDskin, provide an encouraging opportunity to derive in vitro quantitative information related to the inherent sensitizing potency of chemicals, which can be used for potency-associated rankings, for GHS potency sub-categorization as a complement to GARDpotency, or for direct incorporation into existing strategies for QRAs.

Methods

Chemicals. The chemicals used in this study were purchased from Sigma Aldrich (St Louis, Missouri), and their identities are listed in Table 1. They were selected to include a majority of skin sensitizers, which would allow for evaluation of the hypothesis that cDV_0 was associated with potency. A set of non-sensitizers was also included to verify that they did not give rise to positive responses over the examined concentration ranges. The skin sensitizers were selected to form a set of chemicals with varying skin sensitizing potencies, ranging from compounds generally considered as weak sensitizers to those considered as relatively stronger.

Cellular exposure experiments. The experimental procedures were adapted from the default protocols of the GARDskin assay. The GARDskin method has previously been described in detail and the complete protocol has been made publicly available[24]. The scientific validity and performance of GARDskin has also recently been confirmed, in a detail peer review[41]. Nevertheless, briefly described, the GARDskin assay is initiated by evaluating the cytotoxic properties of a test chemical by exposing the cell line, Senzacell (ATCC depository PTA-123875), to the chemical at a range of concentrations. The cell line is a human myeloid dendritic-like cell line, acting as a surrogate for dendritic cells. It is a stable cell line derived from MUTZ-3 (available from DSMZ, ACC 295) by adaptation to simplified growth conditions and cell maintenance protocols (WO 2019/057977), as further described in Ref.[22]. Following 24 h of chemical exposure, cells are harvested and cell viability is examined using propidium iodide staining and flow cytometry analysis. A chemical concentration capable of inducing low- to non-cytotoxic levels is then identified, where low cytotoxicity is defined as a decrease of approximately 10% in cell viability compared to unstimulated controls. This concentration is henceforth referred to as the GARD input concentration. The cytotoxicity screening ensures that chemical effects on the cells are not driven by toxicity, while also homogenizing exposure conditions for all test chemicals. Three independent exposure experiments are then performed at the GARD input concentration. Following 24 h of chemical exposure in each of the experiments, cells are harvested and RNA is isolated and quality controlled, generating three biological replicate samples. The expression levels of the genes in the GARDskin prediction signature (GPS[22,42]) are quantified using NanoString nCounter technology[43], and the skin sensitizing hazard property of the chemical is predicted using the GARDskin prediction model. The exact identity of the genes in the GPS are available in previous publications[22].

For the purpose of dose–response analysis in GARDskin, minor extensions to the original protocols were introduced. For this study specifically, all examined chemicals had previously been investigated in GARD-associated research and validation studies and formerly established GARD input concentrations were used[23,24,42]. Chemical-specific concentrations were selected as described in the section below. Exposure experiments in accordance with the standard GARDskin protocols were carried out at all selected concentrations. However, in contrast to the original protocols, to facilitate the acquisition and enable dose-dependent analysis, each concentration was assayed in one biological replicate. Remaining experimental steps were carried out in accordance with the default GARDskin protocol[24].

Concentration selection. Individual concentration steps were selected for all chemicals. Each set of concentrations was comprised of a geometric sequence with a scaling factor of 3/5. For each chemical, twelve different concentrations were initially considered during the cellular exposure experiments, where the highest concentration corresponded to two steps above the GARD input concentration or to 500 µM, which is in accordance with the standard GARDskin upper concentration limit. The general formula for calculating the different concentration levels is described in Eq. (1). Based on observed cytotoxicity following exposure experiments, eight to nine concentrations per chemical were selected for quantification and classification using the GARDskin pipeline, where concentrations inducing excessive cytotoxicity and/or the lowest evaluated concentration where excluded. The selected concentrations for all chemicals are described in Supplementary Information, and the utilized GARD input concentrations are described in Table 1.

$$c_i = c_{input} \times \left(\tfrac{3}{5}\right)^{i-3},\ i = \{1, 2, \ldots, 12\}, \tag{1}$$

where c_i represent the concentration at step i, and c_{input} the GARD input concentration.

Dose–response analysis. Response values were defined as the decision values generated by the standard GARDskin prediction model, a Support Vector Machine (SVM)[44] that was trained and frozen during assay development, as described[42].

Dose–response relationships were modelled using log-logistic relationships, as defined by Eq. (2), where parameters c and d represent the lower and the upper response limits, respectively, b the slope, and e the effective dose ED50. Additional constraints were also enforced in the fitting procedure. The lower asymptote, c, was kept fixed at the response value of unstimulated controls. Further, the slope parameter, b, was only allowed to take on negative values, thereby generating monotonically increasing curves. Models were fitted using the R-package drc[45].

$$f(x, (b, c, d, e)) = c + \frac{d-c}{1+exp^{b(\log(x)-\log(e))}}. \tag{2}$$

For each chemical, the presence of concentration-dependent response values were evaluated by comparing the fitted log-logistic model against a simpler model represented by the best fitted horizontal line (i.e., the expected shape of the data if the concentration would have no effect on the response values). The models' relative ability

to explain the data was tested using likelihood ratio tests, where a significant result indicated that the more complex model explained the data better. This test is referred to as a no-effect test in the remainder of the text.

The lowest concentration where a positive decision value would be generated, i.e., the cDV_0 concentration, was estimated using two different approaches. In the first approach, cDV_0 was calculated using the log-logistic models. From these models, the uncertainty of an estimate was also evaluated by approximating the values' 95% confidence intervals using inverse regression[46–48]. The second approach for cDV_0 estimation consisted of linear interpolation, which was performed between two points observed on adjacent sides of the decision border, see Eq. (3) below.

$$cDV_0 = c_- - \frac{DV_- * (c_+ - c_-)}{DV_+ - DV_-} \,. \tag{3}$$

where, c_- and c_+ represent the concentration of the negative and the positive point respectively, and DV_- and DV_+ the decision value of the negative and the positive point, respectively.

Comparison with established potency measures. The hypothesis that cDV_0 values were informative of skin sensitizing potency was tested by comparing the estimated cDV_0 concentrations to already established potency measures, including LLNA EC3 values and NOEL values from human repeated insult patch tests (HRIPT). LLNA EC3 concentrations were generally collected from Ref.[32]. However, the EC3 value for methylisothiazolinone was obtained from Ref.[28] (due to a previous incorrect reporting of LLNA-data), an EC3 value of 2.1% was used for diethyl maleate as argued in Ref.[29], and the EC3 value of *trans*-anethol was obtained from Ref.[30]. Human NOEL values were collected from Refs.[33] and [31]. CLP-categories were generated from the EC3-values by applying a 2% classification threshold to discriminate between sensitizers of sub-categories 1A and 1B, as defined by the UN GHS. The strength and significance of the association between cDV_0 and EC3 values, and cDV_0 and human NOEL were evaluated using Pearson's correlation coefficient on log-transformed measures.

Visualization. Figures were created using R v4.0.2[49] with the package ggplot2 v3.3.2[50]. To visualize the acquired dose–response data as shown in Fig. 3, generalized additive models (GAMs) were fitted to the individual chemicals' dose–response curves using the R-package mgcv v1.8-31[51].

Data availability

All data supporting the results are available in the manuscript or supporting materials.

Received: 4 June 2021; Accepted: 3 September 2021
Published online: 23 September 2021

References

1. Kimber, I., Basketter, D. A., Gerberick, G. F., Ryan, C. A. & Dearman, R. J. Chemical allergy: Translating biology into hazard characterization. *Toxicol. Sci.* **120**(Suppl 1), S238–S268. https://doi.org/10.1093/toxsci/kfq346 (2011).
2. Ainscough, J. S., Frank Gerberick, G., Dearman, R. J. & Kimber, I. Danger, intracellular signaling, and the orchestration of dendritic cell function in skin sensitization. *J. Immunotoxicol.* **10**, 223–234. https://doi.org/10.3109/1547691x.2012.711782 (2013).
3. Martin, S. F. *et al.* Mechanisms of chemical-induced innate immunity in allergic contact dermatitis. *Allergy* **66**, 1152–1163. https://doi.org/10.1111/j.1398-9995.2011.02652.x (2011).
4. Martin, S. F. New concepts in cutaneous allergy. *Contact Dermat.* **72**, 2–10. https://doi.org/10.1111/cod.12311 (2015).
5. OECD. *The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins*. https://doi.org/10.1787/9789264221444-en (2014).
6. OECD. *Test No. 406: Skin Sensitisation*. https://doi.org/10.1787/9789264070660-en (1992).
7. Magnusson, B. & Kligman, A. M. The identification of contact allergens by animal assay. The guinea pig maximization test. *J. Investig Dermatol.* **52**, 268–276. https://doi.org/10.1038/jid.1969.42 (1969).
8. Buehler, E. V. Delayed contact hypersensitivity in the guinea pig. *Arch. Dermatol.* **91**, 171–177. https://doi.org/10.1001/archderm.1965.01600080079017 (1965).
9. Ritz, H. & Buehler, E. Procedure for conducting the guinea pig assay. In *Current Concepts in Dermatology* (Academic Press, 1980).
10. OECD. *Test guideline No. 429: Skin Sensitisation*. https://doi.org/10.1787/9789264071100-en (2010).
11. European Chemicals Agency. *Guidance on Information Requirements and Chemical Safety Assessment: chapter R.7a: Endpoint Specific Guidance*. (Publications Office, 2017).
12. Daniel, A. B. *et al.* International regulatory requirements for skin sensitization testing. *Regul. Toxicol. Pharmacol. RTP* **95**, 52–65. https://doi.org/10.1016/j.yrtph.2018.03.003 (2018).
13. Official Journal of the European Union. *OJ L 104, 20.4.2017* 8–11 (Official Journal of the European Union, 2017).
14. OECD. *Test guideline No. 442C: In Chemico Skin Sensitisation*. https://doi.org/10.1787/9789264229709-en (2020).
15. OECD. *Test guideline No. 442D: In Vitro Skin Sensitisation*. https://doi.org/10.1787/9789264229822-en (2018).
16. OECD. *Test guideline No. 442E: In Vitro Skin Sensitisation*. https://doi.org/10.1787/9789264264359-en (2018).
17. Kleinstreuer, N. C. *et al.* Non-animal methods to predict skin sensitization (II): An assessment of defined approaches**. *Crit. Rev. Toxicol.* **48**, 359–374. https://doi.org/10.1080/10408444.2018.1429386 (2018).
18. Natsch, A., Gfeller, H., Rothaupt, M. & Ellis, G. Utility and limitations of a peptide reactivity assay to predict fragrance allergens in vitro. *Toxicol. In Vitro* **21**, 1220–1226. https://doi.org/10.1016/j.tiv.2007.03.016 (2007).
19. Wareing, B. *et al.* The kinetic direct peptide reactivity assay (kDPRA): Intra- and inter-laboratory reproducibility in a seven-laboratory ring trial. *Altex* **37**, 639–651. https://doi.org/10.14573/altex.2004291 (2020).
20. Wareing, B. *et al.* Prediction of skin sensitization potency sub-categories using peptide reactivity data. *Toxicol. In Vitro* **45**, 134–145. https://doi.org/10.1016/j.tiv.2017.08.015 (2017).
21. Natsch, A., Haupt, T., Wareing, B., Landsiedel, R. & Kolle, S. N. Predictivity of the kinetic direct peptide reactivity assay (kDPRA) for sensitizer potency assessment and GHS subclassification. *Altex* **37**, 652–664. https://doi.org/10.14573/altex.2004292 (2020).
22. Johansson, H., Lindstedt, M., Albrekt, A.-S. & Borrebaeck, C. A. A genomic biomarker signature can predict skin sensitizers using a cell-based in vitro alternative to animal tests. *BMC Genom.* **12**, 1–19. https://doi.org/10.1186/1471-2164-12-399 (2011).
23. Zeller, K. S. *et al.* The GARD platform for potency assessment of skin sensitizing chemicals. *Altex* **34**, 539–559. https://doi.org/10.14573/altex.1701101 (2017).

24. Johansson, H. *et al.* Validation of the GARD™ skin assay for assessment of chemical skin sensitizers: Ring trial results of predictive performance and reproducibility. *Toxicol. Sci.* **170**, 374–381. https://doi.org/10.1093/toxsci/kfz108 (2019).
25. Gradin, R. *et al.* The GARDpotency assay for potency-associated subclassification of chemical skin sensitizers-rationale, method development, and ring trial results of predictive performance and reproducibility. *Toxicol. Sci.* **176**, 423–432. https://doi.org/10.1093/toxsci/kfaa068 (2020).
26. Holland-Letz, T. & Kopp-Schneider, A. Optimal experimental designs for dose-response studies with continuous endpoints. *Arch. Toxicol.* **89**, 2059–2068. https://doi.org/10.1007/s00204-014-1335-2 (2015).
27. Slob, W., Moerbeek, M., Rauniomaa, E. & Piersma, A. H. A statistical evaluation of toxicity study designs for the estimation of the benchmark dose in continuous endpoints. *Toxicol. Sci.* **84**, 167–185. https://doi.org/10.1093/toxsci/kfi004 (2005).
28. Roberts, D. W. Methylisothiazolinone is categorised as a strong sensitiser in the Murine Local Lymph Node Assay. *Contact Dermat.* **69**, 261–262. https://doi.org/10.1111/cod.12167 (2013).
29. Roberts, D. W. Interpretation of murine local lymph node assay (LLNA) data for skin sensitization: Overload effects, danger signals and chemistry-based read-across. *Curr. Res. Toxicol.* https://doi.org/10.1016/j.crtox.2021.01.004 (2021).
30. Natsch, A. & Emter, R. Skin sensitizers induce antioxidant response element dependent genes: Application to the in vitro testing of the sensitization potential of chemicals. *Toxicol. Sci.* **102**, 110–119. https://doi.org/10.1093/toxsci/kfm259 (2007).
31. Api, A. M., Parakhia, R., O'Brien, D. & Basketter, D. A. Fragrances categorized according to relative human skin sensitization potency. *Dermatitis* **28**, 299–307. https://doi.org/10.1097/der.0000000000000304 (2017).
32. Hoffmann, S. *et al.* Non-animal methods to predict skin sensitization (I): The Cosmetics Europe database*. *Crit. Rev. Toxicol.* **48**, 344–358. https://doi.org/10.1080/10408444.2018.1429385 (2018).
33. Basketter, D. A. *et al.* Categorization of chemicals according to their relative human skin sensitizing potency. *Dermatitis* **25**, 11–21. https://doi.org/10.1097/der.0000000000000003 (2014).
34. Kimber, I., Gerberick, G. F. & Basketter, D. A. Thresholds in contact sensitization: Theoretical and practical considerations. *Food Chem. Toxicol.* **37**, 553–560. https://doi.org/10.1016/s0278-6915(99)00048-4 (1999).
35. Api, A. M. *et al.* 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 (2008).
36. Griem, P., Goebel, C. & Scheffler, H. Proposal for a risk assessment methodology for skin sensitization based on sensitization potency data. *Regul. Toxicol. Pharmacol.* **38**, 269–290. https://doi.org/10.1016/j.yrtph.2003.07.001 (2003).
37. Jowsey, I. R. Proactive surveillance of contact allergies: An important component of the risk management strategy for skin sensitizers. *Contact Dermat.* **56**, 305–310. https://doi.org/10.1111/j.1600-0536.2006.01082.x (2007).
38. Thyssen, J. P., Johansen, J. D. & Menné, T. Contact allergy epidemics and their controls. *Contact Dermat.* **56**, 185–195. https://doi.org/10.1111/j.1600-0536.2006.01058.x (2007).
39. Patlewicz, G. *et al.* Can currently available non-animal methods detect pre and pro-haptens relevant for skin sensitization?. *Regul. Toxicol. Pharmacol.* **82**, 147–155. https://doi.org/10.1016/j.yrtph.2016.08.007 (2016).
40. Bruning, O. *et al.* A range finding protocol to support design for transcriptomics experimentation: Examples of in-vitro and in-vivo murine UV exposure. *PLoS One* **9**, e97089. https://doi.org/10.1371/journal.pone.0097089 (2014).
41. European Commission. Joint Research Centre. *ESAC Opinion on the Scientific Validity of the GARDskin and GARDpotency Test Methods*. (Publications Office, 2021). https://doi.org/10.2760/626728.
42. Forreryd, A., Zeller, K. S., Lindberg, T., Johansson, H. & Lindstedt, M. From genome-wide arrays to tailor-made biomarker readout—Progress towards routine analysis of skin sensitizing chemicals with GARD. *Toxicol. In Vitro* **37**, 178–188. https://doi.org/10.1016/j.tiv.2016.09.013 (2016).
43. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* **26**, 317–325. https://doi.org/10.1038/nbt1385 (2008).
44. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297. https://doi.org/10.1023/A:1022627411411 (1995).
45. Ritz, C., Baty, F., Streibig, J. C. & Gerhard, D. Dose-response analysis using R. *PLoS One* **10**, e0146021. https://doi.org/10.1371/journal.pone.0146021 (2016).
46. Draper, N. R. & Smith, H. *Applied Regression Analysis* (Wiley, 1998).
47. Greenwell, B., Kabban, C. & Schubert,. investr: An R package for inverse estimation. *R J.* **6**, 90–100 (2014).
48. Schwenke, J. R. & Milliken, G. A. On the calibration problem extended to nonlinear models. *Biometrics* **47**, 563–574. https://doi.org/10.2307/2532146 (1991).
49. R Core Team. (R Foundation for Statistical Computing, 2020).
50. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
51. Wood, S. N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B* **73**, 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x (2011).

Acknowledgements
The authors would like to thank Dr. Rikard Alm for project assistance, and Dr David W. Roberts for critically reviewing the manuscript.

Author contributions
R.G., A.F., and H.J. conceptualized and designed the study and wrote the manuscript. R.G. performed all data analysis and prepared figures. U.M. and A.J. performed the experiments that generated the data. All authors reviewed and approved the manuscript.

Competing interests
SenzaGen AB funded the study in its entirety. All authors are employees of SenzaGen AB.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-98247-7.

Correspondence and requests for materials should be addressed to H.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021

Supplementary materials

Assayed concentrations

Supplementary table 1. Assayed concentrations and response values for the different chemicals are described.

Substance	Assayed concentration	Decision value
1-Butanol	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	−2.67, −2.37, −0.786, −2.39, −1.08, −1.11, −1.33, −1.35
2-Hydroxyethyl acrylate	167, 100, 60, 36, 21.6, 13, 7.78, 4.67, 2.8	10.5, 4.12, 2.48, −0.0909, −0.852, −1.56, −1.44, 0.334, 0.388
2,4-Dinitrochlorobenzene	8.3, 5, 3, 1.8, 1.1, 0.65, 0.39, 0.23, 0.14	2.77, 3.7, 2.33, −1.11, −1.09, −1.9, −1.19, −2.32, −2.14
3-Dimethylaminopropylamine	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	0.752, 0.3, −0.44, −0.576, −1.17, −0.959, −0.521, −0.41
7-Hydroxycitronellal	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	3.2, 3.37, 3.33, 2.46, 2.96, 0.104, −0.175, −0.39
Anethole	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	−1.11, −1.38, −0.938, −2.52, −2.39, −1.13, −1.36, −1.5
Benzalkonium chloride	3, 1.8, 1.1, 0.65, 0.39, 0.23, 0.14, 0.084, 0.05	3.75, 2.58, 0.7, −0.448, −1.26, −1.31, −1.09, −1.72, −1.12
Benzocaine	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	−0.323, −0.894, −0.772, −1.35, −0.945, −0.961, −1.68, −0.888
Benzyl Alcohol	500, 300, 180, 108, 64.8, 38.9, 23.3, 14, 8.4, 5.04	0.909, 2.14, 0.437, 1.07, 0.768, 0.717, 1.02, 0.271, −0.846, −0.727
Benzyl salicylate	334, 200, 120, 72.1, 43.2, 25.9, 15.6, 9.34, 5.6	0.00677, 0.687, −0.829, −1.47, −2.34, −2.32, −2.46, −1.73, −1.35
Cinnamic aldehyde	60, 36, 21.6, 13, 7.8, 4.7, 2.8, 1.7, 1	13.8, 18.3, 11, 11.7, 5.83, 0.421, −0.671, −1.39, −2.85
Diethyl maleate	167, 100, 60, 36, 21.6, 13, 7.78, 4.67, 2.8	9.22, 6.49, 4.54, 3.59, 1.93, 0.882, 0.644, 0.389, −2.09
Dimethyl fumarate	90, 54, 32, 19, 12, 7, 4.1, 2.5, 1.5	10.9, 5.9, 4.85, 3.02, 0.114, 0.503, −1.06, −0.878, −1.52
Eugenol	500, 300, 180, 108, 64.8, 38.9, 23.3, 14, 8.4	4.3, 2.21, 1.11, 0.536, 0.158, −0.341, −0.469, −0.487, −0.883
Farnesol	108, 64.8, 38.9, 23.3, 14, 8.4, 5.04	2.61, 0.64, −0.955, −0.44, −1.57, −0.665, −1.43
Geraniol	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	3.29, 1.23, 1.34, 0.589, −0.412, −0.436, −1.49, −0.936
Glycerol/Glycerin	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	−1.12, −1.98, −1.49, −1.38, −1.71, −1.2, −1.05, −1.75
Imidazolidinyl urea	50, 30, 18, 10.8, 6.5, 3.9, 2.3, 1.4, 0.84	1.4, −1.04, −1.9, −1.04, −1.63, −0.823, −1.6, −1.18, −1.45
Iodopropynyl butylcarbamate	16.7, 10, 6, 3.6, 2.16, 1.3, 0.78, 0.47, 0.28	3.28, 0.777, 0.152, −1.26, −1.41, −1.53, −1.48, −1.13, −1.08
Isoeugenol	500, 300, 180, 108, 64.8, 38.9, 23.3, 14, 8.39, 5.04	5.98, 2.69, 0.386, 0.595, 0.619, 0.463, 0.232, 0.52, −0.284, −0.03

Kanamycin sulfate	208, 125, 75, 45, 27, 16.2, 9.7, 5.8, 3.5	−2, −2.66, −1.69, −2.69, −1.65, −1.7, −1.11, −1.36, −1.06
Linalool	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	0.162, −0.62, 0.0923, −0.752, −0.464, −1.09, −1.34, −1.1
Methylisothiazolinone	16.7, 10, 6, 3.6, 2.2, 1.3, 0.78, 0.47, 0.28	2.35, 0.969, −0.834, −0.665, −0.178, −1.22, −0.635, −1.66, −2.3
Octanoic acid	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	−0.105, −0.515, −1.25, −1.98, −1.48, −1.64, −2.57, −2.01
Pentachlorophenol	150, 90.1, 54, 32.4, 19.5, 11.7, 7, 4.2, 2.5	4.22, 0.75, −1.09, −1.58, −0.493, −0.579, −1.78, −1.61, −1.9
Phenol	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	1.21, −0.114, −0.931, −0.925, −0.275, −1.88, −1.57, −1.31
Salicylic acid	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	−0.807, −0.863, −1.11, −1.05, −0.93, −0.808, −0.892, −0.61
Vanillin	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	−0.605, −0.829, −0.876, −1.04, −0.528, −0.838, −0.948, −1.4
Xylene	500, 300, 180, 108, 64.8, 38.9, 23.3, 14	−0.816, −0.73, −0.996, −0.494, −1.27, −1.3, −1.12, −0.572

Comparison of potency performance between assays

Supplementary Table 2. Overlapping set of chemicals when comparing GARDskin Dose-Response's estimated cDV_0 values to kDPRA's log k_{max} and hCLAT's minimum induction threshold (MIT).

Substance	CAS	Mw (g/mol)	LLNA EC3 (%)	LLNA pEC3 (M^{-1})	GARDskin Dose-Response cDV0 (mg/L)	hCLAT MIT (mg/L)	kDPRA log k_{max} (k_{max} in units of $s^{-1}M^{-1}$)
2,4-Dinitrochlorobenzene	97-00-7	202.55	0.06	337	0.443	1.92	−0.56
3-Dimethylaminopropylamine	109-55-7	102.18	2.2	4.64	25.7	82.1	NA
7-Hydroxycitronellal	107-75-5	172.26	22.2	0.776	5.70	26.8	−2.77
Cinnamic aldehyde	104-55-2	132.16	1.15	11.5	0.524	10.2	−1.35
Diethyl maleate	141-05-9	172.18	2.1	8.20	0.754	64	−1.21
Dimethyl fumarate	624-49-7	144.13	0.35	41.2	0.874	3.94	−0.27
Eugenol	97-53-0	164.21	12.9	1.27	9.29	64.4	−2.64
Farnesol	4602-84-0	222.37	4.8	4.63	12.1	21	−3.41
Geraniol	106-24-1	154.25	23.2	0.665	12.7	123.03	−3.41
Imidazolidinyl urea	39236-46-9	388.29	24	1.62	14.9	39.3	−1.11
Iodopropynyl butylcarbamate	55406-53-6	281.09	0.9	31.2	1.61	8.15	0.75
Linalool	78-70-6	154.25	30.4	0.507	43.0	68.3	NA
Methylisothiazolinone	2682-20-4	115.15	0.4	28.8	0.904	7.89	−0.25
Pentachlorophenol	87-86-5	266.34	20	1.33	20.1	19.04	NA
Isoeugenol	97-54-1	164.21	1.35	12.2	1.70	NA	−1.24

Supplementary Table 3. Correlation coefficient comparing the different assay's potency measure to LLNA on overlapping set of chemicals. Set 1 corresponds to chemicals with data in both GARDskin Dose-Response and hCLAT, set 2 to chemicals with data in both GARDskin Dose-Response and kDPRA, and set 3 to chemicals with data in all three assays.

	Measure	**Pearson correlation**	**p-value correlation**	**N**
Overlapping set 1	GARDskin cDV_0	0.823	2.97×10^{-4}	14
	hCLAT MIT	0.799	6.07×10^{-4}	14
Overlapping set 2	GARDskin cDV_0	0.875	1.95×10^{-4}	12
	kDPRA log k_{max}	0.753*	4.74×10^{-3}	12
Overlapping set 3	GARDskin cDV_0	0.874	4.35×10^{-4}	11
	hCLAT MIT	0.882	3.32×10^{-4}	11
	kDPRA log k_{max}	0.752*	7.63×10^{-3}	11

* The correlation for kDPRA's log kmax with LLNA was based on pEC3 values (as described in [1]), because the assay's potency measure is in units of $s^{-1}M^{-1}$.

hCLAT MIT values were extracted from the Cosmetics Europe database [2].

kDPRA log k_{max} values were obtained from [1].

1 Natsch, A., Haupt, T., Wareing, B., Landsiedel, R. & Kolle, S. N. Predictivity of the kinetic direct peptide reactivity assay (kDPRA) for sensitizer potency assessment and GHS subclassification. *Altex* **37**, 652-664, doi:https://doi.org/10.14573/altex.2004292 (2020).

2 Hoffmann, S. *et al.* Non-animal methods to predict skin sensitization (I): the Cosmetics Europe database<sup/>. *Crit Rev Toxicol* **48**, 344-358, doi:https://doi.org/10.1080/10408444.2018.1429385 (2018).