Research Article

The GARD™ skin Assay: Investigation of the Applicability Domain for Metals

Andy Forreryd¹, Robin Gradin¹, Olivia Larne¹, Nissanka Rajapakse², Eliot Deag² and Henrik Johansson³

¹SenzaGen AB, Lund, Sweden; ²Johnson Matthey, Royston, UK

Abstract

New approach methods (NAMs) for hazard identification of skin sensitizing chemicals have been adopted as test guidelines by the OECD during the last decade as alternatives to animal models. These models align to individual key events (KE) in the adverse outcome pathway (AOP) for skin sensitization for which the molecular initiating event (MIE) is covalent binding to proteins. As it currently stands, the AOP does not include mechanistic events of sensitization by metals, and limited information is available on whether NAMs accurately predict sensitization potential of such molecules, which have been proposed to act via alternative mechanisms to organic chemicals. Methods for assessing the sensitization potential of metals would comprise valuable tools to support risk management within e.g., occupational settings during production of new metal salts or within the medical device industry to evaluate leachables from metal alloys. This paper describes a systematic evaluation of the applicability domain of the GARD™ skin assay for assessment of metals. Hazard classifications were supplemented with an extended analysis of gene expression profiles induced by metal sensitizers to compare the induction of toxicity pathways between metals and organic sensitizers. Based on the results of this study, the accuracy, sensitivity, and specificity of GARD™ skin for prediction of skin sensitizing hazard were 92% (12/13), 100% (7/7) and 83% (5/6), respectively. Thus, the performance of GARD™ skin for assessment of metals was found to be similar to what is observed on conventional organic substances, providing support for inclusion of metals within the applicability domain of the test method.

1 Introduction

Skin sensitizers are compounds that possess the intrinsic potential to induce an immunological hypersensitivity reaction in humans, that upon repeated topical exposure may result in the development of allergic contact dermatitis (ACD). The molecular mechanisms of skin sensitization induced by low molecular weight (LMW) chemicals have been reviewed in numerous publications (Kimber et al., 2011; Martin, 2015), and have also been summarized in an Adverse Outcome Pathway (AOP) by the OECD (OECD, 2014). Of particular importance, ACD develops in two phases, involving initial asymptomatic immunological priming of antigen specific CD4+ Th cells and cytotoxic CD8+ T-cells resulting in the generation of an adaptive immunological memory, which upon subsequent re-exposure to the same antigen will give rise to a rapid clonal expansion and proliferation of effector cells responsible for driving the adverse skin reaction at the site of exposure (Kimber et al., 2011). Thus, in contrast to the reversible damage associated with skin irritation, which is independent of adaptive immune activation, acquired sensitization is generally irreversible, and may result in elicitation of clinical symptoms upon each subsequent exposure, and thus remains a common consumer and occupational health problem. In this context, proactive identification and evaluation of skin sensitizing potential are of central importance for safety evaluation of chemicals within the occupational setting and represents a key toxicological endpoint among regulatory authorities across multiple industries (Daniel et al., 2018, Strickland et al., 2019).

Toxicological hazard assessments for the endpoint of skin sensitization have seen a fundamental change in direction during the last two decades, aiming to replace traditional animal models, such as the guinea pig-based assays (OECD TG 406) (OECD, 2021a) and the murine-based LLNA assay (OECD TG 429) (OECD, 2010) with mechanistically based and scientifically sound in vitro, in chemico and in silico assays, collectively referred to as New Approach Methods (NAMs). To date, a total of nine such methods, which align to the individual key events (KE) of the AOP for skin sensitization (OECD, 2014), have been formally validated in ring-trials, independently reviewed, and incorporated by OECD under the KE-specific Test Guideline (TG) numbers 442C (OECD, 2021b), D (OECD, 2018a) and E (OECD, 2022), respectively. To replace animal
models in regulatory settings, the readout from several NAMs targeting different KE in the AOP needs to be combined using either a weight-of-evidence (WoE) approach, or any of the Defined Approaches (DA) specified in the OECD TG 497 (OECD, 2021c), where the latter is independent of expert judgement and uses a fixed Data Interpretation Procedure (DIP) to convert the readout from multiple information sources into a final classification.

The development, validation, and inclusion of NAM-based strategies, and later also DAs, into the OECD TGs has been an important milestone for replacing animal models and to generate the necessary trust among end-users that results will be accepted by relevant authorities. However, it is relevant to consider that the majority of all assays, both animal and non-animal based, were initially developed, and validated using chemicals from a rather narrow subset of the potentially infinite chemical space (see for example the validation study reports for DPRA (EURL ECVAM, 2011) and h-Clat (EURL ECVAM, 2012). To better understand the applicability domain of individual methods, empirical data for diverse chemical classes must be generated to better understand limitations and provide guidance to end-users for the selection of the most appropriate method(s) depending on test chemical chemistry.

Of particular note, a chemical space for which there to date exists little information in the scientific literature regarding the applicability of NAMs to predict the skin sensitization potential is inorganic molecules, such as metals. The mechanisms underlying sensitization towards metals are generally not as well understood as compared with those of organic compounds. The major immunological signals are similar and converge at T-cell activation, which occurs via interactions between T-cell receptors (TCR) and major histocompatibility complexes (MHC) in combination with the relay of secondary co-stimulatory signals (Kimber et al., 2011). However, while organic compounds form antigenic molecules by binding covalently to endogenous proteins, metals are believed to act via alternative processes which may be protein-independent (Riedel et al., 2021). For example, metals are not expected to form covalent bonds to proteins, and some compounds have been shown to possess the ability to circumvent the classical antigen processing pathways by interacting directly with peptide/MHC complexes or with TCR – peptide/MHC interfaces. The incomplete understanding of the sensitization mechanisms likely impedes inference regarding NAM’s applicability to accurately assess such compounds. However, the limited information available regarding NAM’s applicability for metals may be concerning, as a variety of metallic elements, including e.g., nickel (Schuttelhaar et al., 2018), cobalt (Makrilas et al., 2010), and palladium (Faurschou et al., 2011), have been associated with the potential to induce allergic reactions in humans. As metals may exist in a variety of different salts, organometallic compounds, and alloys intended for use within medical devices (Eichenbaum et al. 2021, Uter et al., 2020), both the potential chemical space and the risk from human exposure may be considerably underestimated.

In this work, the KE3-based Genomic Allergen Rapid Detection™ (GARD™skin) assay (Johansson et al., 2011), which was recently adopted into OECD TG 442 E as the first harmonised assay that generates and interprets genomic data for a regulatory endpoint, was utilized for assessing the skin sensitization properties of metals. For this purpose, a total of 13 compounds, including a variety of metal species and salt forms, predominantly with existing information on skin sensitization potential from human experience and/or animal testing, were evaluated. The results from this study suggest that GARD™skin may contribute to expanding the current applicability domains of NAM-based approaches to also cover metals and thus expand the toxicologist’s toolbox with a non-animal-based approach capable of delivering reliable results for test materials associated with this chemical space.

2 Materials and methods

2.1 Chemicals

The following compounds were evaluated in the GARD™skin assay: cisplatin, nickel (II) sulphate hexahydrate, palladium di(4-oxoent-2-en-2-oate), hydrogen hexahydroxy platinate, (trans) diamminedichloropalladium, cobalt chloride, potassium dichromate, potassium permanganate, zinc sulphate, all of which were obtained from Sigma Aldrich (St Louis, Missouri), and diammonium hexachloroplatinate, tetrammine palladium (II) hydrogen carbonate, tetrammineplatinum (II) hydrogen carbonate and a proprietary platinum salt, all of which were obtained from Johnson Matthey (London, UK). The proprietary platinum salt will remain coded and referred to as JM proprietary Pt salt throughout this manuscript. Table 1 provides additional details on the chemicals, including CAS number, molecular weight, linear formula, purity, and oxidation state.

2.2 GARD™skin assay protocol and experimental conditions

All testing was performed at SenzaGen’s GLP-compliant laboratory (Lund, Sweden) and performed under GLP-like conditions. Experimental procedures were conducted according to the GARD™skin assay protocol (EURL ECVAM, 2021) and in compliance with the GARD™skin method of the OECD TG 442E for testing of single substances (OECD, 2022), with a minor adaption considering the selection of vehicle for the cellular exposure experiments. For the purpose of this study, DMSO was not considered an appropriate vehicle due to concerns regarding the specific solvation chemistry, particularly with platinum group metal compounds. Platinum has a high affinity for sulfur, and when dissolved in DMSO, it will result in ligand displacement and changes to the structure of the complex, potentially interfering with the toxicity profile of the parent molecule (Hall et al., 2014; IPA, 2017). Instead, alternative (inorganic) solvents, such as Dulbecco’s Phosphate Buffered Saline (DPBS, Cytiva), distilled water and cell media (MEM-alpha, Active) were prioritized and evaluated. The solubility of each material was initially assessed by preparing a 10x stock-solution in cell media, or a 1000x stock solution in water or DPBS, and then further diluted in cell media into a default maximum in-well concentration of 500 μM. For test chemicals not soluble in any of the above solvents, less polar solvents not considered to have an impact on the chemical speciation, such as dimethyl formamide (DMF), were explored. These were prepared by first preparing a 1000x stock solution, prior to dilution to final in-well concentration of 0.1%. Table 2 provides a summary of the selected vehicles for each chemical in the study. The experimental vehicles were included as additional negative controls at corresponding in-well concentrations. Following the selection of appropriate vehicle, experimental procedures were performed as described in the GARD™skin assay protocol (EURL
ECVAM, 2021). In short, the human myeloid dendritic-like cell line, SenzaCells™ (available from DSMZ, ACC 295) was exposed to a test chemical at a single concentration, referred to as the GARD input concentration, which was established based on the solubility or cytotoxicity profiles of the individual chemicals. Cells were exposed in three individual experiments in order to generate three independent biological replicate samples for each test chemical. Following 24h incubation, cells were harvested and RNA isolated and quality controlled. Transcriptomic level of the genes in the GPS, whose identity have been transparently disclosed in several publications (see for example Forreryd et al. 2016) was used to assign a Decision Value (DV) to each individual biological replicate of each test chemical. Final assignment of a test chemical as a skin sensitizer or non-sensitizer was strictly based on the mean DV from the biological replicate measurements, were a mean DV < 0 is classified as a non-sensitizer (UN GHS no category), and a mean DV ≥ 0 is classified as a skin sensitizer (UN GHS Category 1), without acknowledging borderline ranges, in full compliance with the GARDskin protocol applied during the formal validation study.

### 2.3 Exploratory data analysis

The induced gene expression profiles were explored using differential expression analysis, visualization of pairwise associations, and principal component analysis (PCA). Differential expression analysis was performed using limma version 3.52.0 (Ritchie et al., 2015; Phipson et al., 2016), which is a statistical framework for differential gene expression analysis based on linear regression models. It uses the empirical Bayes method for borrowing information across genes, and to model the effects induced by metal sensitizers and organic sensitizers (each group were compared with non-sensitizers; 7 metal sensitizers, 18 organic sensitizers, and 21 non-sensitizers). Genes identified with a false discovery rate below 0.05 were considered significantly differentially expressed between compared conditions.

For the correlation-based analysis, chemical specific log fold changes were extracted from the differential expression analysis and pairwise Spearman correlation coefficients were estimated between every compound. The correlation coefficients were visualized using a heatmap created using the R-package heateRmap, version 1.0.12, where the visualized dendrograms were created using complete-linkage clustering on Euclidean distances. Finally, the PCA was created on centered and scaled (mean = 0, standard deviation = 1) log fold changes. Matrix factorization was performed using singular value decomposition, and the principal components were calculated by multiplying the standardized variables by the right singular vectors.

**Tab. 1: Test chemical information and descriptors**

<table>
<thead>
<tr>
<th>Chemical</th>
<th>CAS #</th>
<th>Molecular weight (g/mol)</th>
<th>Linear formula</th>
<th>Oxidation state</th>
<th>Purity</th>
<th>Impurities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisplatin</td>
<td>15663-27-1</td>
<td>300.05</td>
<td>Pt(NH₃)₂Cl₂</td>
<td>II</td>
<td>≥ 99%</td>
<td>No information available</td>
</tr>
<tr>
<td>Cobalt chloride</td>
<td>7646-79-9</td>
<td>129.84</td>
<td>CoCl₂</td>
<td>II</td>
<td>97%</td>
<td>No information available</td>
</tr>
<tr>
<td>(trans)-Diaminedichloropalladium</td>
<td>13782-33-3</td>
<td>211.39</td>
<td>Pd(NH₃)₂Cl₂</td>
<td>II</td>
<td>≥ 99%</td>
<td>No information available</td>
</tr>
<tr>
<td>Diammonium hexachloroplatinate</td>
<td>16919-58-7</td>
<td>443.87</td>
<td>(NH₄)₂PtCl₆</td>
<td>IV</td>
<td>≥ 99%</td>
<td>No information available</td>
</tr>
<tr>
<td>Hydrogen hexahydroro platinate</td>
<td>51850-20-5</td>
<td>299.14</td>
<td>H₃Pt(OH)₆</td>
<td>IV</td>
<td>≥ 99%</td>
<td>Total metallic impurities: 96 ppm</td>
</tr>
<tr>
<td>JM proprietary Pt salt</td>
<td>Confidential</td>
<td>Confidential</td>
<td>Confidential</td>
<td>Confidential</td>
<td>≥ 99%</td>
<td>No information available</td>
</tr>
<tr>
<td>Nickel (II) sulphate hexahydrate</td>
<td>10101-97-0</td>
<td>262.85</td>
<td>NiO₂S.6H₂O</td>
<td>II</td>
<td>≥ 98%</td>
<td>Total metallic impurities: 150 ppm</td>
</tr>
<tr>
<td>Palladium (II) oxide-2-oate</td>
<td>14024-61-4</td>
<td>304.64</td>
<td>C₁₀H₃O₂Pd</td>
<td>II</td>
<td>≥ 99%</td>
<td>No information available</td>
</tr>
<tr>
<td>Potassium dichromate</td>
<td>7778-50-9</td>
<td>294.18</td>
<td>K₂Cr₂O₇</td>
<td>VI</td>
<td>≥ 99%</td>
<td>No information available</td>
</tr>
<tr>
<td>Potassium permanganate</td>
<td>7722-64-7</td>
<td>158.03</td>
<td>KMnO₄</td>
<td>VII</td>
<td>≥ 99%</td>
<td>No information available</td>
</tr>
<tr>
<td>Tetraammine palladium (II)</td>
<td>134620-00-1</td>
<td>296.58</td>
<td>C₂H₇N₂O₅Pd</td>
<td>II</td>
<td>≥ 99%</td>
<td>Total metallic impurities: 76 ppm</td>
</tr>
<tr>
<td>hydrogen carbonate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tetraammine platinum (II)</td>
<td>123439-82-7</td>
<td>385.23</td>
<td>C₂H₇N₂O₅Pt</td>
<td>II</td>
<td>≥ 99%</td>
<td>Total metallic impurities: 40 ppm</td>
</tr>
<tr>
<td>hydrogen carbonate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zinc sulphate</td>
<td>7733-02-0</td>
<td>161.45</td>
<td>ZnSO₄</td>
<td>II</td>
<td>No information available</td>
<td></td>
</tr>
</tbody>
</table>
In the absence of protein reactivity reported in a study using a slightly modified protocol of the DPRA assay (Hemming et al., 2012), Cobalt chloride is reported as a skin sensitizer in humans (Bauch et al., 2012).

4 Data from animal models and human experience available in the ECHA registration dossier suggest that diammonium hexachloroplatinate is a skin and respiratory sensitizer (ECHA, 2018). This is in line with available evidence indicating that chloroplatinates, such as hexachloroplatinate, are potent sensitizers in humans (Clear et al., 1976). Nickel (II) sulphate hydrate is a frequent sensitizer in humans and included in the European baseline series (EBS) of contact allergens, see for example (Schutte et al., 2018, Uter et al., 2020).

5 No human data available. Classified as a skin sensitizer in LLNA study performed in accordance with OECD TG 429 (ECHA, 2017a). No EC3 was calculated. Initial data were indicative of strong sensitizer. Full LLNA was not performed.

6 Classified as a skin sensitizer in LLNA (Basketter & Scholes, 1992) and a skin sensitizer in humans (human NOEL 111, Basketter et al., 2014).

7 No human data available. Results from the guinea pig maximization test (GPMT), conducted to GLP and according to OECD Test Guideline 406. Tetraamine palladium (II) hydrogen carbonate was classified as a skin sensitizer (ECHA, 2017b).

8 No human data available. Classified as a non-sensitizer in an LLNA study when tested up to 25%, conducted according to OECD Test Guideline 429 study and to GLP (ECHA, 2017c).

9 JM proprietary Pt salt is classified as a non-sensitizer in an LLNA study performed in accordance with OECD TG 429 (personal communication with JM).

10 Classified as a non-sensitizer based on data from a GLP compliant GPMT study performed in compliance with OECD TG 406 as reported in (ECHA, 2010), induction 1% intradermal injection, induction 10% topical application, and 0.1% for challenge.

11 Classified as a non-sensitizer based on literature assessment and absence of protein reactivity reported in a study using a slightly modified protocol of the DPRA assay (Hemming et al., 2019). Classified as a non-sensitizer in Buher and GPMT (ECHA, 2017e); challenged with 50% following a 2-week induction period involving three 6 h epicutaneous occlusive applications at 50%; challenged with 75% following a two-stage induction with 25% by intradermal injection and 75% by topical application.

12 Classified as a non-sensitizer in LLNA (Basketter et al., 1999). Tested up to 25%.

3 Results

3.1 Assay compatibility for testing of metals

A pre-validation study was performed prior to initiation of the main GARD™skin study to evaluate the solubility and cytotoxicity profiles of the metal salts to ensure compatibility of the test materials with assay components. Two materials, palladium di(4-oxopent-2-en-2-oate) and hydrogen hexahydroxy platinumate, presented some challenges relating to solubility and were insoluble in the stock solutions in the preferred vehicles in this study (cell media and water). In an attempt to further increase final in-well concentrations for these materials, alternative vehicles were explored. Based on the results from the extended solubility testing, DMF was deemed the most appropriate solvent for palladium di(4-oxopent-2-en-2-oate), resulting in a maximum in-well concentration of 60 μM. In contrast, hydrogen hexahydroxy platinumate was insoluble in the majority of evaluated solvents at the maximum tested concentration of 500 μM, including ethanol, acetone and DMF (data not shown).

Instead, a less concentrated stock solution of 10mM in DPBS was prepared, rendering a stable dispersion, which could further be diluted in cell media to a final in-well concentration of the test material of 100 μM (10% w/v DBPS). For all remaining test materials, no solubility issues were reported using the solvents indicated in Table 2, and they were fully soluble into a maximum in-well concentration of 500μM, thereby adhering to standard GARD™skin protocols.

Following the solubility testing, cytotoxicity profiles were generated by examining the relative viability of cells exposed to the test materials within a titrated range of different concentrations, starting from the default maximum concentration of 500 μM, or the highest soluble concentration. Results are illustrated in Figure 1, and the derived RV90 concentrations for
the cytotoxic materials are summarized in Table 2. As shown in the figure, five of the materials, including JM proprietary Pt salt 1, tetraammineplatinum (II) hydrogen carbonate, hydrogen hexahyroxylplatinate, potassium permanganate and zinc sulphate, did not induce sufficient cytotoxicity at tested concentrations to derive an RV90 concentration. Four of the test materials, including diammonium hexachloroplatinate, cisplatin, palladium di(4-oxopent-2-en-2-olate), and potassium dichromate were highly cytotoxic, rendering RV90 concentrations of 80 µM, 7 µM, 2 µM, and 1.5 µM respectively. Remaining test materials, including tetracmine palladium (II) hydrogen carbonate, nickel (II) sulphate hexahaldrate, (trans) diamminedichloropalladium, and cobalt chloride, demonstrated low cytotoxicity, rendering RV90 values between 450 - 500 µM.

Overall, based on the results from the pre-validation study, all test materials, except for hydrogen hexahyroxylplatinate, were either freely soluble at the maximum default concentration, or induced cytotoxicity at lower concentrations, complying with all acceptance criteria specified in the GARD™skin assay protocol. For hydrogen hexahyroxylplatinate, the material was soluble at a maximum in-well concentration of 100 µM and did not induce cytotoxicity at any of the assessed concentrations. Nevertheless, previous studies in GARD™skin have demonstrated that most sensitizers, irrespective of their sensitizing potential, are detected at concentrations below 100 µM (Gradin et al., 2021). Therefore, despite a lower solubility and no apparent cytotoxicity, hydrogen hexahyroxylplatinate was also considered as acceptable. To conclude, based on the above presented results, evaluated metals were considered as compatible with the already established chemical preparation and cellular exposure protocols for organic chemicals, and did not impose any apparent issues from a technical perspective.

3.2 GARD™skin classifications

Following the completion of the pre-validation study, which resulted in the selection of appropriate vehicles and the establishment of material-specific GARD input concentrations, downstream testing in the GARD™skin assay was performed in full compliance with established protocols, including cellular exposure experiments, preparation of RNA material, gene expression analysis, and data processing using the established analysis pipeline, with no further amendments or modifications. The resulting DVs from the SVM classifications of the individual biological replicates are visualized in Figure 2. Final classifications of test materials as skin sensitizers and non-sensitizers are based on the mean DV from the biological replicates, using a fixed classification threshold (Table 2, mean DV ≥ 0 for a skin sensitizer classification).

3.3 Comparison of GARD™skin classifications to available reference data

To determine the predictive value of GARD™skin for the subset of investigated metals, the binary hazard classifications from the assay were compared with existing reference data on skin sensitization potential, predominantly based on human experience, results from animal testing, or a combination. The concordance between GARD™skin classifications and reference data are summarized in Table 3, and further detailed for the individual test materials in Table 2, together with justifications and sources for the reference classifications. In total, based on the 13 metal containing compounds evaluated in this study, the accuracy, sensitivity, and specificity of GARD™skin for prediction of skin sensitizing hazard were 92% (12/13), 100% (7/7) and 83% (5/6), when compared with available reference data. The 95% confidence interval (95% CI) for the accuracy in this study, calculated with the ClopperPearson method, was (64.0%, 99.8%). Further, the p-value for evaluating the null hypothesis of an accuracy of 50% (H0: 0.5) was 0.0034, demonstrating statistically significant results.
Fig. 2: GARDskin classifications
Summary of DVs generated for the individual biological replicates (n=3) for each test material. A mean DV (indicated by a horizontal line) ≥ 0 is the classification threshold for a test item to be classified as a skin sensitizer.

<table>
<thead>
<tr>
<th>Reference Classifications</th>
<th>GARD™skin Classifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>S</td>
<td>S</td>
</tr>
</tbody>
</table>

Accuracy: 92.3%
Sensitivity: 100%
Specificity: 83.3%

A single non-concordant result was reported for the test material (trans) diamminedichloropalladium, which was classified as false positive in GARD™skin when compared with Local Lymph Node Assay (LLNA) data, as reported in an EU REACH registration dossier where it was tested at a maximum concentration of 25%. Furthermore, for the JM proprietary Pt salt, GARD™skin results are indicative of a borderline classification, with DVs on both sides of the binary classification threshold (DV: -1.8, 0.01, 0.59), albeit final classification (based on mean DV) were negative. Available LLNA data (unpublished) for this compound was difficult to interpret, as a decrease in response with increasing dose over the whole dose range tested was observed. However, this Pt salt was found to elicit a cytokine fingerprinting profile consistent with respiratory sensitizing activity (selective Th2-type cytokine secretion) corresponding to that described previously as associated with a number of respiratory allergens (Dearman and Kimber, 2001). The only definitive conclusion to be drawn from the LLNA was that JM proprietary Pt salt is not a potent contact sensitiser in this test system, which corresponds well with the absence of allergic dermatitis cases in humans. Thus, on the basis of the weight of evidence, this substance was not classified as a skin sensitizer (JM internal data). The negative, albeit borderline, classification in GARD™skin aligns with this conclusion.

Based on available data, it was concluded that GARD™skin provided highly concordant classifications to the reference data, supporting the inclusion of metals into the applicability domain of the method.

3.4 Exploratory data analysis and mapping of toxicity pathways activated by metals
The transcripts in the biomarker signature have been associated with a variety of pathways of relevance to the sensitization process, such as innate immune recognition, oxidative stress, and dendritic cell activation (Johansson et al., 2011), and are thus “in line with mechanisms described under key events of the skin sensitization AOP” (Corsini et al., 2021). To gain further insight into possible similarities and differences in molecular mechanisms and toxicity pathways induced by metal sensitizers in comparison to organic sensitizers, the gene expression profiles of the subset of metals evaluated in this study were compared...
Fig. 3: Correlation heatmap comparing gene expression profiles induced by the metals in this study to historical GARDskin training data

The heatmap displays Spearman correlation coefficients for pairwise comparisons between induced log fold changes of the genes in the GPS (n=196) for evaluated metals and historical data from the GARDskin training dataset, which comprises mainly low molecular weight (LMW) organic chemicals (18 sensitizers and 20 non-sensitizers). Materials are labeled based on dataset (metals in green, training set in orange) and sensitizing potential according to reference data (non-sensitizer in green, sensitizer in orange). The color of a tile reflects the Spearman correlation coefficient. The order of the samples was based on hierarchical clustering, as described by the dendrogram.

With gene expression profiles from the training dataset of GARD™skin, comprising mainly low molecular weight (LMW) organic chemicals (18 sensitizers, 20 non-sensitizers). Figure 3 shows a heatmap of Spearman correlation coefficients of pairwise comparisons between log fold changes for the metals investigated in this study compared to unstimulated control samples, side-by-side with historical data from the GARD™skin training dataset. As expected, overall highest correlations were observed among samples with similar GARD™skin classifications, i.e., samples classified as either skin sensitizers or non-sensitizers, indicating a high similarity in the gene expression profiles driving classification outcomes. For the metals, some of the test materials induced very similar gene expression profiles. Of special interest, (trans) diaminedichloropalladium, which was classified as a GARD™skin false positive in this study, demonstrated an overall high similarity to the sensitizers in the dataset, particularly to the tetraammine palladium (II) hydrogen carbonate (Spearman = 0.93). Moreover, among the non-sensitizers, weak metal-specific correlation structures associated with the platinum species could also be observed, as evident by the slightly higher spearman correlation coefficients for the pairwise comparisons of the JM proprietary Pt salt, hydrogen hexahydroxy platinate, and tetraammineplatinum (II) hydrogen carbonate, compared to unstimulated control.

Overall, comparing the induced gene expression profiles of the metals classified as skin sensitizers and the GARD™skin training data, no unique expression structures associated with the metals could be observed, indicating an overall similarity in the pathways driving the classifications as skin sensitizers for the metals and the organic chemicals for the investigated dataset. Interestingly, however, among the classified sensitizers, the metal sensitizers demonstrated highest concordance with the extreme or strong organic sensitizers in the GARD™skin training dataset.

To further investigate the gene expression induced by the metals in comparison with organic chemicals, a PCA analysis was performed. Figure 4A and B illustrates a PCA plot of the two first components for the materials in the GARD™skin training data and the metals analyzed in this study, using the gene expression from the GPS as variable input.
Fig. 4: Gene expression induced by metals compared to low molecular weight organic chemicals

Principal Component Analysis (PCA) were used to compare gene expression profiles induced by the subset of metals to historical gene expression profiles from the training dataset of GARDskin, comprising mainly low molecular weight (LMW) organic chemicals (18 sensitizers and 20 non-sensitizers), using the genes in the GPS (n=196) as variable input. Materials are colored according to sensitizing potential based on reference classifications (Sensitizers in orange, non-sensitizers in green). The metals evaluated in this study are encircled and labeled. (A) The Euclidean space has been zoomed to facilitate the visualization of the training data, excluding the test material cobalt chloride, which demonstrated a unique expression profile compared to other materials. (B) Visualization of the complete Euclidean space, including the test material cobalt chloride.

(n=196). The metal sensitizer cobalt chloride gave rise to particularly strong effects on the genes, as evident by its distant position compared with other compounds in the figure. For the remaining compounds, a separation between sensitizers and non-sensitizers was observed along the first and the second component for both the training data and the metals. Metals classified as non-sensitizers clearly overlap with non-sensitizers in the training dataset, while metals classified as skin sensitizers, with the exception of the above-mentioned test material cobalt chloride, occupied a similar space as the training dataset, again demonstrating a high similarity in the gene expression profiles of metals and organic chemicals in this dataset.

Finally, a more detailed analysis of genes responsible for driving classifications as skin sensitizers were performed by comparing the subset of differential expressed genes (DEGs) for the metal sensitizers and the organic sensitizers, respectively, versus non-sensitizers. Based on this analysis, no genes with completely different expression patterns were identified (i.e., up and down regulations). Furthermore, genes and signaling pathways responsible for driving the positive classifications of metal sensitizers overlapped with those previously demonstrated to be activated by organic chemical sensitizers, including Nuclear factor erythroid 2-related factor (Nrf2) pathway activation (HMOX1, NQO1, TXNRD1, GSR, SLC5A6, RXRA, MGST3), upregulation of co-stimulatory molecules in DCs (CD86), regulation of toll-like receptor signaling pathway (TLR6, LPY96, CD86, MAP2K1, MAPK13, TLR9), and activation of the intracellular proinflammatory signaling pathways, such as the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB). A list of differentially expressed genes for the investigated metal sensitizers is available in Table S1. Based on the above presented data, metals and organic chemical allergens appear to share similar activation pathways in the herein investigated DC-like cell model.

4 Discussion

During the last decade, several mechanistically based NAMs targeting individual KE in the skin sensitization AOP have been developed and incorporated into OECD TG 442 C, D, and E, for the prediction and classification of chemical skin sensitizers. Results from these assays, when combined into a weight-of-evidence approach or a DA, are capable of overcoming some of the limitations associated with the individual assays and have been demonstrated to provide a similar level of information or be more informative than the LLNA assay. The OECD TG 497 describes DAs for binary hazard classification and potency classification. They are deemed suitable for replacing currently used in vivo methods within a variety of chemical sectors and geographical regions (for an overview of the regulatory landscape, please see Daniel et al., 2018 and Strickland et al., 2019). However, due to properties inherent to a specific test system, or the biomarkers governing the classifications, not all assays may be equally applicable for a specific chemistry. Thus, to provide confidence in classification outcomes, careful characterization of the applicability domain of a given method is critical.

While certain subsets of chemical space, such as hydrophobic substances and indirectly acting haptons, have been recognized to be difficult to accurately assess in one or several NAM-based methods (Mehling et al., 2019, Bergal et al., 2020), the lack of systematic evaluations and the limited availability of data for the testing of metals, metal salts, and organometallic compounds, has made it difficult to arrive at a solid conclusion whether such compounds would fall within the applicability of these methods.

1 doi:10.14573/altex.2203021s
The lack of data for metals may partly be explained by the perceived inability of the current NAM-based strategies to accurately assess such compounds, since metals “are known to react with proteins with mechanisms other than covalent binding” (OECD, 2014), and thus do not perfectly align to the current AOP-based testing paradigm for testing of organic molecules, for which the molecular initiating event (MIE) is covalent binding to endogenous proteins (OECD, 2014; Sullivan et al., 2017). In this regard, the OECD TG 442C clearly states that methods under this TG are not applicable for the testing of metal compounds, since they are known to react with mechanisms other than covalent binding to proteins, while OECD TGs 442D and 442E state that methods have shown to be applicable to test chemicals covering a variety of organic functional groups but provide no specific guidance for metals or other inorganic molecules. In contrast, in vivo models, such as the LLNA assay provide acceptable performance for metals, with the exception of nickel, which is a well-characterized false negative in the assay (Basketter et al., 1999; ICCVAM, 2010).

In the current study, the applicability of the GARD™ skin assay for hazard assessment of metal contact allergens was investigated by comparing binary hazard predictions with existing reference data on skin sensitization potential, predominantly based on human experience, results from animal testing, or a combination. In this context, it should be acknowledged that some commentators have highlighted the importance of applying a so-called triangular approach for establishing the predictive performance of NAM-based methods, including a comparison to both human and animal reference data, as well as a head-to-head comparison between human and animal data (Natsch et al., 2021). This strategy is indeed valid and provides a comprehensive assessment of the predictivity of the evaluated method, while accounting for potential conflicting reference data between humans and animals. However, it is imperative to consider that such a strategy is dependent on the availability of well-curated reference data from both human and animal tests. For example, for ethical reasons, human sensitization data obtained from Human Repeated Insult Patch Testing (HRRIPT) are rarely available for substances not explicitly designed for use in products in direct contact with human skin (e.g., cosmetics). For this reason, human data were available in the literature only for a limited number of the substances in this study (n=5). The limited availability of human reference data is far from unique to the present study, and a recent example includes the assessment of agrochemical formulations (Strickland et al., 2022), for which the authors reported all performance metrics of evaluated NAMs exclusively versus available animal data. In addition, animal data from the LLNA assay or the GPMT were available for a total of eight and three compounds, respectively. While the authors acknowledge that individual comparisons of GARDskin predictions versus animal, or human data would indeed be interesting, the relative scarcity of available reference data from each data source would indeed limit the usefulness of such an approach. Thus, for the purpose of evaluating predictive performance in this manuscript, priority was given to human data when available (only the compound nickel had conflicting animal and human data), and animal data were only used when no human data were available. Importantly, while some recent publications have reported issues related to low reproducibility (Dumont et al., 2016), or high number of false positives (summary available in Natsch et al., 2021) for the LLNA assay, it should be emphasized that this assay is still considered as the method of choice in many regulatory jurisdictions (Daniel et al., 2018) and have also a proven capability to accurately predict the sensitization potential of metals (Basketter et al., 1999, ICCVAM, 2010). Furthermore, it should be noted that significant efforts have been made in order to browse available literature for available references, and the final assignment of sensitization potential, as well as references supporting the classification, is transparently reported for each test compound in Table 2.

Based on the 13 metal compounds evaluated in this study, the accuracy, sensitivity, and specificity of GARD™ skin for prediction of skin sensitizing hazard were 92% (12/13), 100% (7/7) and 83% (5/6), when compared with available reference data. Thus, the overall predictive performance of GARD™ skin for hazard assessment of metals was similar to previously reported estimations of predictive performance for organic low-molecular-weight chemicals, providing support for the inclusion of metals into the applicability domain of the test method. The reported performance metrics for GARD™ skin can be considered in the context of results from other predictive tests. For the LLNA assay, a systematic evaluation of 13 metal salts demonstrated a balanced accuracy of 87%, with a false negative and a false positive classification reported for nickel chloride and copper chloride, respectively (Basketter et al., 1999). Furthermore, also zinc sulphate, which was correctly labeled as a non-sensitizer in this study has been reported as a false positive in the LLNA (ICCVAM, 2010). Nevertheless, there seem to be some discrepancies in reported results for this metal, since other LLNA studies have correctly classified this compound as a non-sensitizer (Basketter et al., 1999). Similar systematic evaluations on the performance of the validated NAMs in OECD TG 442 for metals are essentially lacking in the scientific literature, with the exception of a recent case study describing the use of the DPRA assay, slightly modified, for assessment of a variety of platinum and palladium salts (Hemming et al., 2019). While this modified version of the DPRA is formally not compliant with OECD TG 442C in terms of regulatory acceptance, results from this study are encouraging, but it remains to be evaluated how the assay performs across a wider subset of metals. Moreover, data from other OECD validated NAMs are available only for a very limited subset of metals. For example, in the comprehensive database presented by Urbisch and colleagues (Urbisch et al., 2015), which contains results for a total of 213 substances, only five were categorized as inorganic salts. As evident from this publication, some metals, such as potassium dichromate and cobalt chloride were correctly labeled in the investigated test guideline assays, while others, such as for example nickel chloride, zinc sulphate and beryllium sulphate induced mixed results. For example, nickel chloride was correctly classified as a skin sensitizer in h-CLAT but rendered a false negative classification in LuSens (Urbisch et al., 2015), and induced a mix of positive and negative classifications in KeratinoSens (EURL ECVAM, 2011). Furthermore, the non-sensitizer zinc sulphate has been reported as a false positive in the h-CLAT assay (Urbisch et al., 2015), while the skin sensitizer beryllium sulphate induced a mix of positive and negative classifications (EURL ECVAM, 2013). Furthermore, a recent publication describes the use of the IL-18 RHE assay, a non-guideline in vitro approach based on reconstructed human epidermis (RHE) for the assessment of a variety of metal salts. In this study, 7/9 metal allergens, including the herein correctly predicted nickel (II) sulphate hexahydrate, failed to induce positive classifications in the assay, resulting in a sensitivity of only 22%, indicating that metals fall outside the applicability domain of this assay (Gibbs et al., 2018). Thus, based on the lack of systematic evaluations, and the inconsistency in the reported results, currently available experimental data are not sufficient to arrive at a firm
conclusion regarding the applicability, or inapplicability domain, of the above-mentioned assays for inorganic molecules or metals.

While the work presented in this paper provides empirical evidence of accurate predictions of the GARD™skin assay for a variety of metals, it is important to proactively consider potential restrictions, if any, which would not support the inclusion of metals into the applicability domain of the GARD™skin in a regulatory context. Firstly, from a technical perspective, the tested metals were directly compatible with assay components, enabling testing in accordance with the validated GARD™skin protocol, with only a minor adaption considering the specific solvent used to dissolve the materials prior to cellular exposures. Precise understanding of metal solubility/helution and bioavailability is key to the potential for metal ions to have sensitizing hazard (IPA, 2017). In this context, it should be stressed that the use of alternative solvents to DMSO in this study was a deliberate choice based on the chemistry of the test materials. Platinum, for example, has a high affinity for sulfur, and when dissolved in DMSO, it will result in ligand displacement and changes to the structure of the complex, potentially interfering with the toxicity profile of the parent molecule (Hall et al., 2014; IPA, 2017). Instead, water and assay media were successfully used to dissolve the test materials to the maximum default concentration for all test materials, except for hydrogen hexahydrated platinate and palladium di/(4-oxopent-2-en-2-oate), where alternative vehicles were required to enhance solubility. Importantly, the highest applied concentration was only limited by solubility for a single test material, hydrogen hexahydrated platinate. However, using an alternative vehicle (DPBS), a final in-well concentration of 100 μM could be obtained, which based on previously reported data is within the concentration range where a signal from a potential sensitizer would be detected in the assay (Gradin et al., 2021). Thus, using a combination of carefully selected standard and alternative vehicles, appropriate measures could be taken not to alter the chemical speciation or the test materials, while increasing the test concentrations above the limit of detection of the assay. In the context of the above discussion, it should be considered that the OECD TGs 442 C, D, and E, support the use of alternative solvents/vehicles to those specified in the guideline when scientifically justified and encourage consideration of chemical stability of the test material in the selected solvent. Thus, the herein reported use of alternative solvents should not per se constitute an obstacle for direct implementation of metals in the applicability domain of GARD™skin.

Secondly, it is relevant also to consider if the subset of metals in this study is sufficiently representative of the metal space. The rationale for the selection of test materials in this study was to include a variety of different metals for which there were available reference data in the literature for evaluation of assay performance. Details on references used for sensitization categorization of each material are summarized in Table 2. Nickel was a special case since it is a frequent cause of ACD in humans but still fails to induce a response in the LLNA assay, and the molecular mechanisms responsible for the observed discrepancy between the animal and human data are discussed in detail in sections below Furthermore, the study also included a variety platinum and palladium salts. These comprised a useful subset of compounds to investigate, since they exhibit a variety of sensitization potentials, including both skin sensitizers and non-sensitizers. Amongst the evaluated platinum compounds, diammonium hexachloroplatinate and cisplatin are recognized as skin sensitizers, based mainly on urticaria in workers and non-contact urticaria in patients, respectively, while hydrogen hexahydrated platinate and JM proprietary Pt salt are considered non-sensitizers. All these compounds were correctly classified by GARD™skin. Similarly, for palladium, tetraamine palladium (II) hydrogen carbonate and palladium di/(4-oxopent-2-en-2-oate) were correctly classified as skin sensitizers by GARD™skin, while the third palladium salt, (trans) diaminedi chloropalladium, was non-concordant with available reference data, indicating a potential false positive response. Reasons for this single discrepancy in the dataset are currently unknown, however differences in dissolution of the metal ion in the GARDskin assay media versus the vehicles used in the LLNA assay may play a part. In summary, the results demonstrate the applicability of the GARD™skin assay to a variety of clinically important metals, such as nickel and cobalt, where the latter is a common constituent in a wide range of medical devices (Eichenbaum et al., 2021). Furthermore, the results also demonstrate that the assay can provide valuable data to support the risk management in the occupational setting during the production of new platinum compounds for use within the electronics, or medical device industry, and during manufacturing of platinum-based catalysts, as well as defining limitations on the product application areas. The number of test materials in the present study is of similar size to the 14 metals that were present in the LLNA database at the time when ICCVAM published the LLNA Applicability Domain Evaluation Report in 2010, resulting in a recommendation to support the inclusion of metals, with the exception of nickel, into the applicability domain of LLNA (ICCVAM, 2010). The reported accuracy from that study, although excluding nickel from the calculations, was lower than the herein reported performance for GARD™skin. Therefore, the predictive performance, as well as the composition and size of the herein reported GARD™skin dataset may be considered as sufficient for inclusion of metals into the applicability domain of the assay. The inclusion of metals into the applicability domain of GARD™skin may thus contribute to reduce the need for animal testing, serving as a scientifically justified stand-alone approach for non-regulatory testing during research projects and product development, but also as an information source within a weight-of-evidence (WoE) approach to comply with regulatory requirements. In such a regulatory context, GARDskin was recently adopted into OECD TG 442E following the positive recommendation from the EURL ECVAM Scientific Advisory Committee (ESAC) (Corsini et al., 2021). Thus, dependent on regulatory context, positive classifications may be used as a stand-alone information source to identify skin sensitizers, while negative result have to be considered together with additional evidence. For example, to comply with information requirements in REACH, the latter can be accomplished either directly by providing supporting data from other KE in the AOP in a WoE approach, or by the subsequent adoption of the assay into any of the DAs described in the OECD TG 497 (ECHA, 2021). Furthermore, in this context, it is also imperative to mention that the current protocol of the GARDskin assay, similar to many other in vitro assays is dependent on animal-derived serum (fetal calf serum, FCS) for cultivation of cells, but work is currently ongoing to adapt the assay to animal component free cultivation protocols (manuscript under preparation).

Thirdly, in a regulatory context, available OECD TGs are KE-based and strictly align to the AOP for skin sensitization, which is currently considered to be applicable only to organic molecules, for which the molecular initiating event (MIE) is covalent binding to endogenous proteins (OECD, 2014). In this context, metals have largely been excluded from the
applicability domain since they are considered to act via alternative mechanisms. For example, following the validation of the DPRA assay, on the basis of assumed understanding of the metal mechanisms, EURL ECVAM concluded that the assay was not designed for the identification of metal allergens (Casati et al., 2013; EUROL ECVAM, 2012), which was later also incorporated in the OECD TG as a known limitation to the assay (OECD, 2021a). Nevertheless, when challenged with a dataset of metals, a slightly modified version of the assay showed promising results (see reference (Hemming et al., 2019)), highlighting that inference on applicability domains should ultimately be based on a scientific method involving empirical data collection and hypothesis testing. Nonetheless, although the AOP, as it currently stands, is not assumed to fully reflect the skin’s immune response towards metals, some mechanisms may still overlap to those initiated by organic chemicals, and thus it may still serve as a viable framework for designing testing strategies. In this context, it is interesting to recapitulate the current science by which sensitizing metals are known to induce allergic hypersensitivity reactions in humans, which has been excellently reviewed in (Riedel et al., 2021).

First, pathomechanism of metal allergy, similar to all adaptive immune responses, are dependent on at least two main signals, including both cell surface antigen presentation by DCs in the context of major histocompatibility complexes (MHC) to naïve allergen specific T-cells, as well as a co-stimulatory signal, or so-called danger signal (Matzinger, 1998; Schmidt et al., 2010). Considering the former, low-molecular weight (LMW) chemicals must first bind to endogenous proteins to generate an immunogenic hapten-protein complex (Landsteiner & Jacobs, 1936). The polarized nature of metals enables the activation of electrons from donor atoms, such as amino acid side chains of appropriate proteins in the skin. However, in contrast to classical haptenes, which binds covalently to proteins, the associated coordinate bond between the metal and the protein need not be irreversible, and it has been hypothesized that such bonds, in contrast to covalent bonds, may not be sufficiently stable to survive the intracellular processing mechanisms required for surface expression on MHC complexes (Chipinda et al., 2011).

Interestingly, although the exact molecular mechanisms of the haptenation remains unknown for most metals, it has been demonstrated that nickel may bypass the conventional antigen processing pathways by direct binding to peptides loaded on the MHC complex from the extracellular space, or to conserved residues at the T-cell receptor - peptide-MHC interface to activate the T-cells in a protein independent manner (Riedel et al., 2021). Furthermore, for nickel, the co-stimulatory signal is delivered by direct interaction between the nickel ion and non-conserved histidine residues on human toll-like receptor 4 (TLR4) on DCs, resulting in the activation of the nuclear factor erythroid 2-related factor 2 (Nrf2) pathway and production of inflammatory cytokines (Schmidt et al., 2010). These non-conserved histidine residues are species-specific and not present on the equivalent receptor in mice, which effectively explains why nickel fail to provoke the necessary inflammatory signals to initiate and sustain an adaptive immune response in mice, and subsequently give rise to a false negative classification in the LLNA assay. Moreover, recent data has also indicated a similar TLR4 dependent activation for cobalt and platinum (Rachmawati et al., 2013). Moreover, although exact mechanisms remains to be explored for palladium, mono-sensitization is uncommon, and hypersensitivity to palladium is often observed together with nickel allergy, indicating cross reactivity, and overlapping mechanisms (Faurschou et al., 2011). Altogether, this is interesting, since it would suggest that sensitization to metals may be acquired also independently on the conventional protein binding mechanisms, and that the direct interaction between metal ions and DCs may be sufficient to activate the inflammatory response driving the induction of the downstream sensitization reactions. In this context, considering the central role for DCs in the activation of the immune sensitization cascades required for sensitization to metals, the relevance of DC based methods for evaluation of the skin sensitization potential of metals is supported by available science on the molecular mechanisms behind metal allergy. This is further substantiated by the GARD™skin results in this study, which suggest that the response patterns for the metal compounds do not differ significantly from the response patterns of organic compounds, potentially allowing for accurate discrimination between sensitizers and non-sensitizers also for metals.

In conclusion, the GARD™skin assay was shown to be applicable for the successful hazard assessment of skin sensitizers in the chemical domain of metals. The evaluated metals were compatible with the assay protocols, and only a minor adaption associated with the selection of vehicle was introduced to ensure sample stability. Furthermore, the predictive performance, the number of tested metals, and the current understanding of the pathomechanisms responsible for metal allergy together provide support for the inclusion of metals into the applicability domain of the GARD™skin assay. Regulatory acceptance and inclusion of metals into the test guideline will be an important step to support risk management within occupational settings and to define limitations on the product application areas for new metal salts, as well as for the medical device industry, significantly reducing the need for animal testing within a variety of industries.

References

11


Conflict of interest
Johnson Matthey authors have no financial interest regarding the GARD™ platform assays, and their development or execution. SenzaGen AB is the method developer of GARD™skin.

Acknowledgements
The authors would like to express thanks to Dr Nicoleta Muresan for providing the test materials sourced from Johnson Matthey.

Data availability
Raw data were generated at the SenzaGen testing facility. Derived data supporting the findings of this study are available from the corresponding author AF on request. The identity of the Proprietary Pt salt will not be disclosed.