Skin Irritation and Sensitization: Mechanisms and New Approaches for Risk Assessment

2. Skin Sensitization

D. Baskettera  R. Darlenski b  J.W. Fluhr c

a DABMEB Consultancy Ltd., Sharnbrook, UK; b Department of Dermatology and Venereology, Medical Faculty, Sofia, Bulgaria; c bioskin, Berlin, Germany

Introduction

In contrast to irritation, skin sensitization is a response of the adaptive immune system, in which there is a delayed T-cell-mediated allergic response to chemically modified skin proteins [1]. In theory, almost any chemical which can covalently modify skin proteins can trigger this type of allergy. Such chemicals are often referred to as haptens or sensitizers. When this type of chemical induces an immune response, an individual may become sensitized, i.e. he/she has a contact allergy, but no disease. However, subsequent skin contact with a sufficient dose of the chemical will elicit a reaction in the skin, called allergic contact dermatitis (ACD). Similar ('cross-reacting') chemicals may also elicit the response. ACD is a relatively common skin disease, being caused by a diverse range of chemicals, including fragrances, preservatives, rubber chemicals, industrial intermediates, epoxy resins, acrylates, medicaments and many others [2].

Attempts have been made in many countries to reduce the problems of ACD by the implementation of legislations related to skin-sensitizing chemicals, as well as by the early detection and risk assessment of substances with sensitizing properties. For many years, the simple identification of sensitizing chemicals was performed in guinea pig tests. A murine test, the local lymph node assay (LLNA), has been validated as a replacement for the guinea pig tests. Despite the recent introduction of in vitro methods for the identification of sensitizing chemicals, the LLNA results (when coupled with good exposure data) can be used as the starting point for a quantitative risk assessment. The quantitative risk assessment is aimed to identify the safe use thresholds for any potential skin sensitizer.

Key Words
Allergic contact dermatitis • Risk assessment • Local lymph node assay

Abstract
Allergic contact dermatitis (ACD) is a common skin disease with a significant social and economic impact. In contrast to irritation, skin sensitization is a response of the adaptive immune system, in which there is a delayed T-cell-mediated allergic response to chemically modified skin proteins. The chemicals that can covalently modify the skin proteins and trigger an allergic reaction are referred to as haptens or sensitizers. Attempts have been made in many countries to reduce the problems of ACD by the implementation of legislations related to skin-sensitizing chemicals, as well as by the early detection and risk assessment of substances with sensitizing properties. For many years, the simple identification of sensitizing chemicals was performed in guinea pig tests. A murine test, the local lymph node assay (LLNA), has been validated as a replacement for the guinea pig tests. Despite the recent introduction of in vitro methods for the identification of sensitizing chemicals, the LLNA results (when coupled with good exposure data) can be used as the starting point for a quantitative risk assessment. The quantitative risk assessment is aimed to identify the safe use thresholds for any potential skin sensitizer.
ally Harmonized System for regulatory toxicology [5]. All these legislations are currently limited in scope to simple binary hazard identification (i.e. whether the chemical is, or is not, a skin sensitizer), although efforts are under way to introduce enhancements which will allow for the discrimination of skin sensitizers of differing potency [6], a topic discussed in more detail below.

The simple identification of chemicals which have the potential to act as skin sensitizers has relied for many years on one of a small number of guinea pig tests (reviewed in Andersen and Maibach [7]). Only two tests are now globally accepted by regulatory authorities, i.e. the Buehler occluded patch test [8] and the guinea pig maximization test [9]. These tests are enshrined in the Organization for Economic Cooperation and Development (OECD) Guideline No. 406 [10]. More recently, a murine test, the local lymph node assay (LLNA), has been formally validated as a complete replacement for the guinea pig tests [11] and has its own test guideline [12]. These tests achieve a similar degree of sensitivity and specificity in terms of hazard identification [13] and thus supply the necessary information for current regulatory toxicology and manufacturers’ safety data sheets. The LLNA, however, has proven to have an advantage over the older guinea pig tests in that it provides a quantitative index of relative sensitizing potency. When coupled with good exposure data, this information can be used as the starting point for a quantitative risk assessment (QRA) which aims to identify safe use thresholds for any potential skin sensitizer [14–17]. The application of risk assessment is much more likely to reduce the extent of clinical disease than the basic hazard identification which feeds the current manufacturers’ safety data sheet system, and accordingly is the main focus of what follows.

**Pathogenetic and Molecular Aspects of Skin Sensitization**

Chemicals that cause ACD do so because of specific properties: they can react (normally covalently) with, and so modify, skin proteins and they can cause at the same time a sufficient degree of skin trauma to engender the production of ‘danger’ signals. To achieve the first of these actions, the substance must gain access to the viable epidermis. Ultimately, the chemically modified protein must be recognized by T cell receptors. Thus, the process of the induction of skin sensitization can be viewed as a series of barriers that a chemical must overcome in order to cause skin sensitization. These have been the subject of extensive reviews recently [1, 18–20]. The main aspects are discussed in more detail below.

The stratum corneum represents a fairly formidable barrier to the entry of noxious materials in healthy skin. However, lipophilic low-molecular-weight (<500 Da) chemicals can pass this barrier relatively easily, a situation likely to be enhanced when skin is damaged or diseased, so that the stratum corneum is no longer fully functional. Skin sensitizers which also cause irritation due to the ability to directly disrupt this barrier may be of particular concern. Prediction of the ultimate epidermal concentration arising from a given degree of skin exposure is currently fraught with difficulty and one must often resort to poor surrogate measures, such as ClogP, the calculated lipophilicity of a chemical [18]. Measures of skin penetration are of limited value, since they indicate what has passed through the skin rather than what has remained there. Having said that, it is reasonable to assume that all chemicals will penetrate the skin to a certain degree, and examples where they fail to sensitize after skin penetration are not commonly known [21].

Assuming that a chemical has penetrated into the viable epidermis, covalent modification of proteins represents a key initiation step for the immune system, but of course the chemical may simultaneously disrupt the cellular function in a direct way, additionally producing a localized trauma signal. The chemistry of protein haptation has been investigated for about 70 years [22, 23], by Mayer in the 1950s [24], by the breakthrough collaboration of Dupuis and Benezra at the beginning of the 1980s [25], and in the 1990s and this century by Lepoittevin and colleagues [26–30]. However, a major step change occurred with the seminal publication of Roberts and Williams in 1982 [31] which combined chemistry with mathematical insights to yield quantitative structure-activity relationships. This led to a whole new body of studies (reviewed in Lepoittevin [19], Basketter et al. [32] and Smith and Hotchkiss [33]), from which most recently has come a new fundamental approach, termed quantitative mechanistic modelling centered on mechanistic applicability domains [34–36]. This work is based on the concept that skin-sensitizing organic chemicals in fact fall within just a few reactivity classes. As this concept is evaluated against existing skin sensitizer databases and those in the process of publication [36], it offers the prospect of a better identification of sensitizing chemicals (even prior to their synthesis), as well as the prediction of potential cross-reactivity. This is of importance, as even the best of the current computer-based predictive systems, DEREK, has limitations [37].

Pathogenetic and Molecular Aspects of Skin Sensitization

Chemicals that cause ACD do so because of specific properties: they can react (normally covalently) with, and so modify, skin proteins and they can cause at the same time a sufficient degree of skin trauma to engender the production of ‘danger’ signals. To achieve the first of these actions, the substance must gain access to the viable epidermis. Ultimately, the chemically modified protein must be recognized by T cell receptors. Thus, the process of the induction of skin sensitization can be viewed as a series of barriers that a chemical must overcome in order to cause skin sensitization. These have been the subject of extensive reviews recently [1, 18–20]. The main aspects are discussed in more detail below.

The stratum corneum represents a fairly formidable barrier to the entry of noxious materials in healthy skin. However, lipophilic low-molecular-weight (<500 Da) chemicals can pass this barrier relatively easily, a situation likely to be enhanced when skin is damaged or diseased, so that the stratum corneum is no longer fully functional. Skin sensitizers which also cause irritation due to the ability to directly disrupt this barrier may be of particular concern. Prediction of the ultimate epidermal concentration arising from a given degree of skin exposure is currently fraught with difficulty and one must often resort to poor surrogate measures, such as ClogP, the calculated lipophilicity of a chemical [18]. Measures of skin penetration are of limited value, since they indicate what has passed through the skin rather than what has remained there. Having said that, it is reasonable to assume that all chemicals will penetrate the skin to a certain degree, and examples where they fail to sensitize after skin penetration are not commonly known [21].

Assuming that a chemical has penetrated into the viable epidermis, covalent modification of proteins represents a key initiation step for the immune system, but of course the chemical may simultaneously disrupt the cellular function in a direct way, additionally producing a localized trauma signal. The chemistry of protein haptation has been investigated for about 70 years [22, 23], by Mayer in the 1950s [24], by the breakthrough collaboration of Dupuis and Benezra at the beginning of the 1980s [25], and in the 1990s and this century by Lepoittevin and colleagues [26–30]. However, a major step change occurred with the seminal publication of Roberts and Williams in 1982 [31] which combined chemistry with mathematical insights to yield quantitative structure-activity relationships. This led to a whole new body of studies (reviewed in Lepoittevin [19], Basketter et al. [32] and Smith and Hotchkiss [33]), from which most recently has come a new fundamental approach, termed quantitative mechanistic modelling centered on mechanistic applicability domains [34–36]. This work is based on the concept that skin-sensitizing organic chemicals in fact fall within just a few reactivity classes. As this concept is evaluated against existing skin sensitizer databases and those in the process of publication [36], it offers the prospect of a better identification of sensitizing chemicals (even prior to their synthesis), as well as the prediction of potential cross-reactivity. This is of importance, as even the best of the current computer-based predictive systems, DEREK, has limitations [37].
Methods and Models for the Evaluation of Skin Sensitization

Guinea Pig Models

The guinea pig has been the animal model of choice for many decades, since the 1930s when Landsteiner and colleagues [38, 39] began basic investigations of the chemistry of skin sensitization. Many of the developed protocols are now history [7]. Two assays, however, are still very much in use, i.e. the Buehler occluded patch test [8] and the guinea pig maximization test [9]. These methods are described in many regulatory guidelines for the assessment of skin sensitization, notably in those of the OECD [10]. It is out of the scope of this study to repeat all the details of the protocols; rather it is sufficient to note that both methods possess a multistep induction phase, followed by a rest period and an elicitation (or ‘challenge’) phase with subjective visual assessment of skin reactions.

When properly conducted, both of these methods provide an adequate degree of sensitivity and specificity for the predictive identification of significant skin sensitization hazards (although this has never formally been validated) [40–42]. However, there is a fair body of literature dealing with interpretive difficulties of these assays, in particular regarding potential weakly positive reactions, for which rechallenge is often recommended. Ultimately, though, skin sensitization testing is moving from the guinea pig to the mouse, a process accelerated by the adoption of the LLNA as the method of first choice in new regulations, notably REACH in Europe. So, most emphasis will be placed on a review of this method and of its strengths and limitations.

The LLNA

Based on a modern appreciation of the immunobiological processes that result in the induction of skin sensitization, opportunities emerged to explore new approaches to skin sensitization testing. The LLNA is predicated upon an alternative strategy in which sensitizing activity is measured as a function of responses induced in mice during the induction, rather than elicitation, phase of contact sensitization [43–45]. In this method, skin sensitizers are identified according to their ability to provoke proliferative responses in draining lymph nodes following repeated topical exposure. In practice, skin-sensitizing chemicals are defined as those which, at one or more test concentrations, induce a 3-fold or greater increase in lymph node cell proliferation compared with concurrent vehicle-treated controls. The LLNA has been the subject of extensive evaluations and the view currently is that the method provides a reliable and robust approach to the identification of sensitizing chemicals and as such represents a stand-alone alternative to guinea pig assays [11, 46, 47]. Following the formal validation of the LLNA [11, 13, 48], it was adopted as the OECD test guideline No. 429 [12], which is a good starting point for finding details of the complete protocol. In brief, the protocol used is typically the following: groups of 4 CBA/Ca female mice (7–12 weeks of age) are treated topically on the dorsum of both ears with 25 μl of test material, or with an equal volume of the vehicle alone (4:1, acetone:olive oil, v/v). Treatment is performed once daily for 3 consecutive days. Five days following the initiation of exposure, all mice are injected via the tail vein with 250 μl of phosphate-buffered saline containing 20 μCi (740 kBq) of tritiated thymidine. The mice are sacrificed 5 h later and the draining lymph nodes excised and pooled for each experimental group. The lymph node cell suspension is washed twice in an excess of phosphate-buffered saline and then precipitated with 5% trichloroacetic acid at 4°C for 18 h. Pellets are resuspended in trichloroacetic acid and the incorporation of tritiated thymidine is measured by liquid scintillation counting.

Results from the LLNA have been published for a wide range of chemicals, including the OECD recommended positive controls [49–52]. After the formal validation of the assay, some groups thought that it may be too sensitive and/or prone to false positives [53, 54]. This topic has been extensively discussed in the literature [41, 54–57]. It is a fact that any toxicology assay has both false positives and false negatives, and the LLNA is no exception. The classic irritant, sodium lauryl sulfate, is a weak, but reproducible positive reactive chemical in the LLNA, but is very widely regarded as lacking significant skin sensitization potential [58, 59]. This substance provides a good benchmark to help distinguish between chemicals which may be true false positives and those which simply are unanticipated (and perhaps unpalatable) true skin sensitizers. It is certainly the case that most of the debate on false positives arose as a result of the crude hazard-based classification systems employed in a number of countries, including those of the EU. However, the greatest strength of the LLNA is that it not only provides hazard identification, but also an assessment of the relative potency, which is discussed in detail below. It should not go unmentioned that potency measures can also contribute to hazard-based regulatory toxicology, an opportunity recommended by many groups, including industry, the EU and the World Health Organization [6, 60].
The LLNA and Measurement of Relative Potency for QRA

In its first form, the LLNA was designed to discover whether a chemical was a potential skin sensitizer. A stimulation index (SI) of 3 or more demonstrated that there was sufficient evidence of the presence of such a hazard. However, no account was taken of whether that simple threshold was exceeded at only a very high concentration, at a low concentration, or somewhere in between. For example, the well-recognized strong sensitizer DNCB gave very high SI values, consistent with this high potency [58]. As experience in the LLNA developed with an increasing range of allergens, it became clear that those generally regarded as weaker, for example in the guinea pig maximization test, gave lower SI values, whereas stronger allergens recorded much higher SI values, particularly when tested at high concentrations [61, 62]. These observations began to stimulate consideration of whether and to what extent LLNA results might reflect not only skin sensitization hazard itself, but also skin sensitization potency. If they did, how might potency be quantified and then be deployed in risk assessment?

Early Investigations

The first real consideration of using LLNA to rank the potency of allergens did not come from comparison with either clinical or earlier in vivo data from the guinea pig, but in fact from a rather different perspective. Attempts had been made for many years to relate chemical structure to skin sensitization effects. Typically, quantitative structure activity relationships (QSARs) were restricted to limited groups of chemicals and the biological data, challenge reactions in guinea pig tests, had to be ‘massaged’ to deliver the type of quantitative readout necessary for QSAR development [63]. However, a young postdoctoral student, Mark Cronin, who had just completed a 3-year study of an extensive guinea pig maximization test database [64], had also spotted that there was an opportunity to use the objective and quantitative LLNA dose response to provide a measure of potency. This was applied to a sensitization data set of a family of bromoalkanes, using manually plotted graphs to determine the intercept of the dose-response curve on the 3-fold threshold line. At that time, the SI value was called the test/control ratio; so that the numbers increased with potency, the reciprocal was used in QSAR development. The publication which arose from this work showed a good graphical correlation and a simple equation was derived to represent the data [65]. The title of the paper ‘The value of the local lymph node assay in structure-activity investigations’ gives an indication of what was beginning to be felt at the time concerning the ease of use of the LLNA and the real opportunity that a relative potency measure such as the reciprocal test/control ratio might represent.

Partly triggered by the above QSAR work, but strongly influenced by the observation that strong sensitizers generated greater SI values than weaker allergens and that this seemed to be in general agreement with guinea pig and clinical evidence, more serious thought was given to how to quantify potency – strictly speaking, relative potency. The first real efforts in this respect are contained in some of the final publications associated with the interlaboratory validation of the LLNA for hazard purposes [58, 66]. In this work, LLNA dose-response curves were fitted using quadratic regression equations, to estimate the concentration at which the curve intersected with the 3-fold threshold line. The outcome of this work summarized in table 1 showed that LLNA potency estimates were rather similar between the 5 laboratories involved (and also served to provide strong evidence of the reproducibility of hazard determination).

In parallel, consideration was given to how this might begin to impact upon risk assessment thinking [62]. In this publication, general issues were aired, but a key point

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Lab A</th>
<th>Lab B</th>
<th>Lab C</th>
<th>Lab D</th>
<th>Lab E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potassium dichromate</td>
<td>0.058</td>
<td>0.132</td>
<td>0.150</td>
<td>0.122</td>
<td>0.126</td>
</tr>
<tr>
<td>Dinitrochlorobenzene</td>
<td>0.027/0.05</td>
<td>0.046/0.06</td>
<td>0.062/0.04</td>
<td>0.094/0.06</td>
<td>0.057/0.03</td>
</tr>
<tr>
<td>Isoeugenol</td>
<td>1.3</td>
<td>3.3</td>
<td>1.8</td>
<td>3.1</td>
<td>1.6</td>
</tr>
<tr>
<td>Eugenol</td>
<td>5.8</td>
<td>14.5</td>
<td>8.9</td>
<td>13.8</td>
<td>6.0</td>
</tr>
<tr>
<td>Hexyl cinnamic aldehyde</td>
<td>7.9</td>
<td>7.6</td>
<td>8.4</td>
<td>7.0</td>
<td>8.1</td>
</tr>
</tbody>
</table>

1 DNCB was tested on two separate occasions.
in the paper was that attention was drawn to the approximately 160-fold difference in potency between DNCB and the weak human allergen hexyl cinnamic aldehyde. What was especially notable was the dynamic range of the measurement. The ratio was calculated, using data from two studies, where the concentration of each chemical necessary to stimulate a 3-fold increase in proliferation in draining lymph nodes compared to concurrent vehicle-treated controls had been derived.

The EC3 Value

On the basis of the thinking outlined above, two specific pieces of work were undertaken. Firstly, a retrospective analysis was carried out to check whether a 3-fold threshold was indeed the most appropriate cutoff for hazard determination in the LLNA [67]. This is discussed in more detail elsewhere, but it is sufficient here to note that the suitability of this limit was confirmed. Secondly, a number of approaches to how to deploy LLNA dose-response data were investigated and the results published [68]. In this work, the clear focus was on the estimation of the concentration of the test chemical necessary to cause a 3-fold stimulation of proliferation, the EC3 value, by interpolation (not extrapolation) of the dose-response data. Initial work evaluated and rejected a variety of complex approaches (this material was not published). Efforts then concentrated on just three methods, the quadratic regression approach already used, a more complex statistical technique using receiver operating characteristics, and linear interpolation between the data points immediately above and below the 3-fold threshold line. Analysis of a range of skin sensitizers demonstrated that all of the methods gave very similar results, certainly within what was already known of the biological variation associated with the assay, and so the simplest method, linear interpolation, was selected as the most appropriate. The details of this method are captured in figure 1. More complex approaches to the derivation of EC3 values were subsequently suggested by others, but they lead to the same numbers as simple linear interpolation [69].

In addition to this, it has long been recognized that there are occasional situations where the LLNA data fail by a small margin to meet the requirements of linear interpolation, but where the data give a sufficient indication of potency, such that additional testing of extra dose groups is not merited. For this reason, details of how to carry out modest degrees of extrapolation to derive EC3 values have been published [70].

Reproducibility of EC3 Measurements

The LLNA is an in vivo biological assay, which means that there must be inherent variability to any measurement. This may reasonably be expected to be compounded further by considerations of time and place (e.g. different laboratories, different staff conducting the assay). In this section, the reproducibility of LLNA EC3 values will be examined in detail. Although some information on reproducibility had already been generated, which was very positive [58], it did not use linear interpolation and it is this standard approach which will be exclusively discussed here.

Table 2 contains the compiled results of testing a number of skin sensitizers on multiple occasions and, where the studies have allowed, the derivation of an EC3 value. This material has been published elsewhere [71], but is presented here in a slightly modified form with a small amount of additional data, notably for potassium dichromate and imidazolidinyl urea. It is instructive to also examine these data in graphical form, presented as the mean and range plus standard error on a log graph (necessary to accommodate the several orders of magnitude that the values span; fig. 2).

Clinical Correlations

As already mentioned, early observations of dose-response curves and of the approximate concentration at which a threshold response was seen had indicated that
### Table 2. Collation of EC3 data from repeat testing of 17 chemicals in multiple laboratories

<table>
<thead>
<tr>
<th>Substance</th>
<th>EC3 values, %</th>
<th>Vehicle</th>
<th>Mean EC3 ± SE¹, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandrowski’s base</td>
<td>0.04, 0.02</td>
<td>AOO</td>
<td>0.03</td>
</tr>
<tr>
<td>2,4-Dinitrochlorobenzene</td>
<td>0.04, 0.02, 0.05, 0.03, 0.03, 0.02, 0.06, 0.03, 0.06, 0.05, 0.05, 0.06, 0.05</td>
<td>AOO</td>
<td>0.04 ± 0.004</td>
</tr>
<tr>
<td>Potassium dichromate</td>
<td>0.05, 0.08, 0.14</td>
<td>DMSO</td>
<td>0.09 ± 0.046</td>
</tr>
<tr>
<td>p-Phenylenediamine</td>
<td>0.07, 0.12, 0.09, 0.08, 0.06, 0.14, 0.06, 0.18, 0.16, 0.13</td>
<td>AOO</td>
<td>0.11 ± 0.014</td>
</tr>
<tr>
<td>1,4-Hydroquinone</td>
<td>0.11, 0.19, 0.12</td>
<td>AOO</td>
<td>0.14 ± 0.04</td>
</tr>
<tr>
<td>Methylidibromoglutaronitrite</td>
<td>1.8, 0.9, 1.3</td>
<td>AOO</td>
<td>1.3 ± 0.45</td>
</tr>
<tr>
<td>Isoeugenol</td>
<td>1.7, 1.1, 1.4, 1.3, 1.0, 1.4, 1.5, 2.9, 0.8, 1.3, 1.6, 2.8, 0.9, 1.0, 1.7, 1.2, 1.4, 0.8, 2.1, 2.3, 1.1, 1.2, 1.2, 0.7, 1.0, 2.3, 1.3, 2.0, 1.6, 1.3</td>
<td>AOO</td>
<td>1.5 ± 0.1</td>
</tr>
<tr>
<td>Cinnamal</td>
<td>3.1, 1.7, 2.7</td>
<td>AOO</td>
<td>2.3 ± 0.4</td>
</tr>
<tr>
<td>1-Bromopentadecane</td>
<td>5.2, 5.1</td>
<td>AOO</td>
<td>5.1 ± 0.02</td>
</tr>
<tr>
<td>L-Perillaldehyde</td>
<td>8.1, 7.8</td>
<td>AOO</td>
<td>8.0</td>
</tr>
<tr>
<td>Hexylcinnamal</td>
<td>6.6, 11.3, 10.6, 4.4, 11.5, 8.8, 7.6, 11.0, 7.0, 10.6, 11.9, 11.7, 10.9, 11.7, 12.2</td>
<td>AOO</td>
<td>9.9 ± 0.6</td>
</tr>
<tr>
<td>Eugenol</td>
<td>15.0, 4.9, 12.9, 7.5</td>
<td>AOO</td>
<td>10.1 ± 2.3</td>
</tr>
<tr>
<td>Abietic acid</td>
<td>14.7, 8.3, 10.6</td>
<td>AOO</td>
<td>11.3 ± 1.8</td>
</tr>
<tr>
<td>Penicillin G</td>
<td>16.7, 17.9, 30</td>
<td>DMSO</td>
<td>21.5 ± 4.3</td>
</tr>
<tr>
<td>Imidazolidinyl urea</td>
<td>23.9, 31.2</td>
<td>DMF</td>
<td>27.6</td>
</tr>
<tr>
<td>Hydroxycitronellal</td>
<td>33.0, 27.5, 23.0</td>
<td>AOO</td>
<td>27.8 ± 2.9</td>
</tr>
<tr>
<td>2-Ethylbutyraldehyde</td>
<td>60, 76</td>
<td>AOO</td>
<td>68</td>
</tr>
</tbody>
</table>

AOO = Acetone:olive oil, 4:1, v/v; DMF = dimethyl formamide; DMSO = dimethylsulfoxide.

¹ Numbers to no more than 2 significant figures; standard error not calculated if there were less than 3 data points.

---

**Fig. 2.** Mean and range graph of EC3 determinations.
such information might correlate well with human potency, but no structured work had been completed. However, before going any further, it is necessary to make one key point: any measure of relative sensitization potency in a predictive assay (in vivo or in vitro) is nothing more than a relative measure of the ability to induce sensitization under defined conditions. Thus, it may be matched against relative human potency for induction of skin sensitization, but it will not be a correlate of the clinical frequency of allergic contact dermatitis any more than it might indicate the likelihood that a particular individual might become sensitized. These are matters of risk assessment, depending heavily on several exposure parameters and individual susceptibility (as well as potency).

Indeed, as might be expected, scant data exist on the relative ability of chemicals to induce skin sensitization in humans. In the first instance, a semi-quantitative approach was taken to the assessment of EC3 values for a range of chemicals with what was understood of their relative human potency. In the primary publication, dermatologists and a regulatory toxicologist were asked to group a set of 20 chemicals into 1 of 5 categories, from potently sensitizing to nonsensitizing. The EC3 values for these chemicals were then mapped onto these categories [61]. The correlation was very good. This work was followed by a similar effort in the USA with 21 chemicals, but using best estimates of thresholds in a predictive human assay as the indicator of relative human potency [72]. The predictive assay, the human repeated insult patch test (HRIPT), had been carried out on a range of chemicals in the latter half of the 20th century. Using a common currency of dose/unit area for both LLNA and HRIPT thresholds, the results again indicated a very good correlation. Of particular note was that relative potency as determined by the EC3 threshold in the LLNA and by the no effect level in the human assays not only agreed quite well, but also spanned a similar range (approx. 4 orders of magnitude).

Subsequently, additional chemicals were progressively added to the list of comparisons of EC3 values with human potency, but the quality of these correlations was limited by the poor availability of human data on many substances. Separately, two groups in Germany, one industry-based and one regulatory group, explored all available human predictive test data, made careful assumptions and then attempted to correlate LLNA potency information with the totality of available human threshold data derived from this human predictive test-
ing [73, 74]. While they contain some differences in approach, both publications clearly show two things – the relatively poor quality of the human data, but notwithstanding, they also indicate that LLNA EC3 values do correlate with indices of human induction potency. In the light of this work, and recognizing the problem of the poor quality of the human predictive test data, a further analysis of the correlation was undertaken, but where the human data were even more carefully critiqued to eliminate poor quality information. The outcome of this assessment with just 25 chemicals again showed a good correlation with LLNA EC3 values [75].

Figure 3 presents the correlation between LLNA EC3 values and all the available human predictive test threshold data from 4 publications [72–75]. The 68 observations show a clear trend, but it is important to remember that while the LLNA EC3 values were obtained under well-controlled conditions, the same is not true for the human data. This no doubt explains part of the discrepancy, but it is also important to recognize that mice would be expected to be a perfect predictor of relative human skin sensitizing potency. Nevertheless, the lines superimposed on the graph at thresholds of 100 μg/cm² (as an example) clearly define a quadrant of the most potent allergens, a potentially useful tool for regulatory toxicology [60, 75].

Most recently, large data sets of EC3 values have been published which extend the opportunity for retrospective analysis of potency comparisons [52]. However, it seems increasingly unlikely (for obvious ethical reasons) that significant quantities of new human potency data will become available. Therefore, the validity of LLNA EC3 values will be determined by their utility in ensuring proper risk assessment and risk management of skin-sensitizing chemicals.

**Quantitative Risk Assessment**

QRA is a relatively new approach to the safety evaluation of skin sensitzers [15, 16]. It arose from the considerations developed with a ‘fragrance think tank’ in 1996/1997, where a considerable effort was being devoted to understand why fragrance allergy was so common. One part of the explanation was that the risk assessment approach deployed was in fact not sufficiently rigorous, which has culminated in a significant strategic change by that industry where QRA has been adopted as the new tool for the establishment of safe limits in a wide range of product types for fragrance allergens [17]. So, what is skin sensitization QRA? An overview of the approach is presented in figure 4. In essence, the potency information from skin sensitization testing, but in particular from the

---

**Fig. 4.** An overview of the QRA approach.
LLNA, is used to predict the threshold for the induction of skin sensitization in an experimental human assay, the HRIPT. This human assay is not carried out, but the indicated threshold (NESIL – see fig. 4) is then adjusted downwards to allow for a number of considerations necessary to render the threshold realistic for the expected exposure to the skin-sensitizing chemical. Thus, an acceptable exposure level (AEL – an upper limit) is defined for a particular sensitizing chemical in a defined use situation. This level can be compared with the consumer exposure level (CEL) which would occur if the product was marketed. If the CEL is above the AEL, then there clearly is a potential problem; if the opposite is the case, then the exposure is likely to be tolerated. The QRA approach can readily be deployed in occupational settings, the key being to ensure that skin exposure is adequately understood.

**In vitro Alternatives**

It would not be appropriate to complete an overview of skin sensitization without making some mention of the strenuous efforts which continue to be directed towards the elimination of the need for the use of animals in the identification and assessment of this type of chemical allergen. There have been a number of recent reviews of all the main elements of this area (chemical, computing and biological), so only key points will be made here. The reader is referred to the reviews themselves for more details [18, 37, 76–81].

Most considerations of nonanimal alternatives for skin sensitization tend to examine a single aspect of the process of induction (e.g. chemical reactivity, epidermal bioavailability, dendritic cell responses). However, experts in the area generally conclude that it will require a combination of data from these endpoints to truly replicate the sensitization process in vitro. Thus, a strategy for the recombination of the various data elements has to be developed as well as the methods themselves. This point is discussed in the work of Jowsey et al. [78], who suggest a holistic approach. That then just leaves the assays themselves to be developed and validated!

The information that indicates that a chemical will be a skin sensitizer is implicit in its molecular structure. Structure activity relationships, both qualitative and quantitative, have been developed, and one, DEREK, regularly evolved to take account of new knowledge [82]. Nevertheless, a recent detailed review suggested that all the available systems were not yet suitable for use as a full replacement for the animal models [37]. A new approach, quantitative mechanistic modeling, holds promise, but awaits more detailed assessment [20].

Chemical reactivity associated with skin sensitization has also been measured via peptide-binding assays [76]. A predictive model of this has recently been developed by Gerberick et al. [83, 84] and has successfully been adopted with only minor modifications by others [85]. It seems quite probable that this area of endeavor will contribute in a significant way in the future [86]. Potentially, a simple combination of such a reactivity measure coupled with information on epidermal bioavailability [18] might provide an adequate basis at least for the predictive identification of skin sensitization hazards.

Ultimately, for a full characterization of potential skin-sensitizing chemicals, it seems most likely that the measures mentioned above will need to be combined with biological data, perhaps from a dendritic cell assay [70, 77, 81]. Indeed, such an assay may provide information on whether a chemical triggers either a specific (allergenic) and/or a nonspecific (irritant) response. Potential approaches to this have been described, but as with the other endpoints mentioned above, all require formal validation [87–89]. Finally, data inputs from a dendritic cell assay would have to be combined with information on peptide reactivity and epidermal bioavailability using the type of paradigm described [78]. The final paradigm itself would depend on the data and would also require some form of validation. It is worth noting that if the chemical reactivity assay provided the necessary specificity of discrimination between sensitizing and irritant properties, then the function of the biological assay might solely be to indicate the extent to which the chemical triggered the production of ‘danger’ signals [90, 91].

Despite the challenges, it has to be said that the prospects for the development of nonanimal alternatives for the skin sensitization endpoint are relatively good. The approaches are well under development and there exists a very solid body of in vivo data against which the new methods can be assessed [52]. Whether both the science and the matter of validation and regulatory acceptance can be achieved within the timescales envisioned by some in Europe remains a matter of some doubt.

**Acknowledgements**

Part of this work was presented at the 3rd International Meeting of the International Society of Skin Pharmacology and Physiology (‘Barrier function of mammalian skin’), Washington, D.C., USA, 2007. R.D. was supported by a scholarship from the German Academic Exchange Service (DAAD), No. A/07/92141, 2007.
References


5 http://www.unece.org/trans/danger/publi/ghs/ghs_rev00/English/GHS-PART-3e.pdf.


24 Mayer RL: Group sensitization to compounds of quinine structure and its biochemical role. The role of these substances in cancer. Prog Allergy 1954;4:79–172.


