Assessing the severity of allergic reactions: a regulatory dilemma

David A. Basketter¹, John P. McFadden² and Ian Kimber³

¹DABMEB Consultancy Ltd, Sharnbrook MK44 1PR, UK, ²Department of Cutaneous Allergy, St John’s Institute of Dermatology, St Thomas’ Hospital, London SE1 EH, UK, and ³Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, UK

doi:10.1111/j.1600-0536.2012.02102.x

Summary

The identification of chemicals possessing the intrinsic ability to cause sensitization via skin contact or inhalation, commonly referred to as skin and respiratory sensitizers, is a key endpoint in regulatory toxicology. Predictive assays for this purpose exist only for skin sensitizers, but, for both types of sensitizer, human evidence can be used to determine whether a substance should be classified. Furthermore, the use of human evidence for subcategorization according to sensitization potency is also accommodated within the regulations. Normally, this is based on the prevalence of sensitization in relation to the degree of exposure in the context of the size of the population exposed. However, the regulations also indicate that the severity of (allergic) reactions may be taken into account. In this article, we consider whether this is appropriate and whether there is evidence that reaction severity can inform decisions on classification and/or potency categorization. The conclusion drawn is that the severity of an allergic reaction does not correlate with, or serve as an indicator of, the sensitizing potency of a chemical. In reality, it reflects the overall extent of sensitization that an individual has acquired, in concert with the concentration of the causative allergen to which they have been exposed.

Key words: globally harmonized system; local lymph node assay; potency categorization; respiratory sensitization; severity; skin sensitization.

It is well established that many chemicals have the potential to cause allergic sensitization of skin and/or of the respiratory tract (1, 2). Proteins are also common causes of respiratory allergy (3), but the current review will focus exclusively on allergy caused by chemicals. If there is sufficient and appropriate exposure, then such sensitization may translate, respectively, into allergic contact dermatitis, or respiratory allergy and asthma.

The predictive identification of chemicals that are able to cause sensitization is a therefore a primary endpoint in regulatory toxicology. Internationally, adoption of the Globally Harmonized System (GHS) of Classification and Labelling of Chemicals, which prescribes, for all toxicity endpoints, the test methods and strategies available for their interpretation continues to evolve (4). For both skin and respiratory sensitization, the GHS text is similar to that of the Dangerous Substances Directive (Directive 67/548/EEC) (5) and the REACH legislation that followed (6). In 2011, the United Nations (UN) issued a fourth revised edition of the GHS, and in the same year the European Commission issued a Second Adaptation to Technical Progress (ATP) implementing the Fourth UN GHS revision (7). The GHS also allows authorities to adopt supplemental labelling provisions to protect individuals who are already sensitized to a specific chemical that may elicit a response at very low concentration. Requirements should be introduced to add the name of such a chemical on the label, even if it is present at a very low concentration in a mixture. However, one key development has been the inclusion of the word ‘severity’ for each sensitization
endpoint. Thus, for the purposes of this review, it is worth noting the specific GHS text from the Fourth revision of the GHS and Second ATP:

<table>
<thead>
<tr>
<th>Category 1</th>
<th>Respiratory sensitizer</th>
</tr>
</thead>
<tbody>
<tr>
<td>A substance is classified as a respiratory sensitizer (1) if there is evidence in humans that the substance can lead to specific respiratory hypersensitivity and/or (2) if there are positive results from an appropriate animal test.</td>
<td></td>
</tr>
</tbody>
</table>

Sub-category 1A: Substances showing a high frequency of occurrence in humans; or a probability of occurrence of a high sensitization rate in humans based on animal or other tests. Severity of reaction may also be considered.

Sub-category 1B: Substances showing a low to moderate frequency of occurrence in humans; or a probability of occurrence of a low to moderate sensitization rate in humans based on animal or other tests. Severity of reaction may also be considered.

<table>
<thead>
<tr>
<th>Category 1</th>
<th>Skin sensitizer</th>
</tr>
</thead>
<tbody>
<tr>
<td>A substance is classified as a skin sensitizer (1) if there is evidence in humans that the substance can lead to sensitization by skin contact in a substantial number of persons, or (2) if there are positive results from an appropriate animal test.</td>
<td></td>
</tr>
</tbody>
</table>

Sub-category 1A: Substances showing a high frequency of occurrence in humans and/or a high potency in animals can be presumed to have the potential to produce significant sensitization in humans. Severity of reaction may also be considered.

Sub-category 1B: Substances showing a low to moderate frequency of occurrence in humans and/or a low to moderate potency in animals can be presumed to have the potential to produce sensitization in humans. Severity of reaction may also be considered.

It should be noted that the word ‘occurrence’ in the above text refers to the presence of induced sensitisation, and not the frequency of exposure.

Materials that have an intrinsic capacity to induce sensitization (of the skin or of the respiratory tract) are defined as representing sensitization ‘hazards’. This indicates that the material possesses this particular intrinsic toxicological property. The acquisition of sensitization as such is not an evident adverse health effect. For the elicitation of an allergic reaction, the sensitized subject has to be exposed subsequently to the same chemical allergen, in sufficient quantity, and via an appropriate route. However, it has long been appreciated that not only do humans vary significantly with respect to their susceptibility to the acquisition of sensitization, but also that chemical allergens themselves show considerable heterogeneity. Most commonly, that heterogeneity is manifest in terms of the relative sensitizing potency, and, indeed, in the case of contact allergens, it is estimated that the relative skin sensitizing potency among contact allergens varies by up to five orders of magnitude (8, 9). Relative potency therefore represents an important, and indeed vital, component of the risk assessment/risk management process (10). Thus, for a potent sensitizer, only relatively low levels of exposure will be required for the acquisition of sensitization, whereas with a weak sensitizer, much higher levels of exposure will be necessary. Naturally, in this context, ‘high’ and ‘low’ are only relative terms that serve to describe the overall extent of exposure with regard to volume, duration, and also frequency. However, the current versions of the GHS documents, in those sections that address allergic sensitization, also note that the severity of reactions may be taken into account, although generally only for the purpose of distinguishing stronger from weaker sensitizers (sub-categories 1A and 1B) (7, 11). However, specifically for skin sensitizers, the concept of severity of reactions is also mentioned in the more general guidance for classification. In the present article, the scientific rationale and practical utility of this is subjected to a critical review.

The mechanistic basis of severity of reactions

From a mechanistic perspective, it is possible to identify three generic factors that will influence the severity of allergic reactions in previously sensitized subjects. The same principles apply both to allergic contact dermatitis and pulmonary allergic reactions.

The first factor is the extent to which a subject is sensitized. In the case of skin sensitization, this equates to the effectiveness and extent of immunological priming, which, in turn, will, at least in part, be a function of the degree to which priming resulted in the clonal expansion of allergen-responsive T lymphocytes. Simplistically, perhaps it can be argued that the larger the pool of allergen-responsive T lymphocytes, the more aggressive will be the secondary response following a subsequent encounter with the same chemical, and the more severe the inflammatory reaction is likely to be. There is less certainty about the effector mechanisms responsible for chemical respiratory allergy, although, in many instances, IgE antibody is believed to play a pivotal role. Irrespective of specific effector mechanisms, the same principle will apply insofar as the degree of immunological priming, and thus the extent to which sensitization has
been acquired, will have an important impact on the severity of pulmonary allergic reactions. In the context of the degree of sensitization achieved, it is important to appreciate that this will be dictated by the conditions (extent, duration, and frequency) of exposure, as well as by interindividual differences in the vigour of induced immune responses. In addition, the inherent potency of the allergen is another important contributor to the level of sensitization that will be achieved. However, in the context of the theme of this article, there is not necessarily a linear relationship between sensitizing potency and the level of sensitization acquired. Thus, prolonged and/or repeated exposure to a weak allergen may result in a level of sensitization equivalent to (or even greater than) that which will result from a single exposure to a low concentration of a strong allergen.

The second factor that will have an important influence on the severity of reactions is the level of exposure of sensitized subjects. It can be argued that, even in subjects with low or moderate levels of sensitization, exposure to a high concentration of the relevant allergen may trigger an aggressive elicitation reaction. Certainly, this is what is observed clinically.

Finally, there are undoubtedly important interindividual differences with regard to the ability to mount inflammatory reactions. One manifestation of this is the well established variation in the response of individuals to non-allergic skin irritants, as exemplified, for instance, by the greater than two orders of magnitude variation in the individual threshold of response to the anionic surfactant sodium lauryl sulfate (12, 13). The consequence is that, for a well defined inflammatory stimulus, individuals will have widely different inflammatory reactions.

Of course, in the real world situation, extrinsic factors and exposure to multiple allergens are critical complicating factors. In relation to the regulatory classification of individual chemicals, the conclusion drawn is that, from a mechanistic perspective, there are a number of factors that may influence the severity of allergic reactions in the skin and lung, and the sensitizing potency of the allergen is only one of these. The corollary is that the severity of allergic reactions does not provide an appropriate basis for considering the relative sensitizing potential of contact and respiratory allergens.

The concept of severity in predictive tests

Predictive tests for respiratory sensitization that are validated and/or accepted in a regulatory setting do not yet exist. Consequently, there is no way in which the topic of severity can be considered in a meaningful way for the predictive identification and consequent regulatory classification of respiratory sensitizers. Furthermore, it has been argued recently that the sub-categorization of respiratory sensitizing chemicals and proteins is premature, not least because of the absence of human benchmark data (14).

For many years, the guinea pig was the species of choice for methods designed for the predictive identification of skin sensitizing chemicals (15). Many protocols were developed, but pre-eminent among them was the guinea pig maximization test (GPMT), which was first described over 40 years ago (16). In this assay, the extent of induced skin sensitization is determined as a function of the frequency of cutaneous reactions resulting from epicutaneous challenge of previously sensitized animals. Grading of the severity of skin responses was viewed as a requirement by the authors of the test, but was not adopted into the description of the potency of a sensitizer (16). However, regulatory guidelines indicate that reaction grades should be recorded on a simple scale that was also mentioned in the monograph:

0 = no visible change
1 = slight or discrete erythema
2 = moderate and confluent erythema
3 = intense erythema and swelling

A similar approach was recommended for the only other regulatory guinea pig assay, the occluded patch test of Buehler (17). Thus, in practice, the severity of individual reactions is recorded in regulatory guinea pig skin sensitization tests. Nevertheless, interpretation of the data, and judgements reached about skin-sensitizing potential, take no account of information about the severity of reactions. The only decision point is the number of guinea pigs that are scored as showing positive skin test reactions – 15% or more in the Buehler test and 30% in the GPMT being the threshold for positive classification (4).

In more investigative studies in the guinea pig, a relationship between the intensity of induction and the elicitation dose threshold has only rarely been reported, but nevertheless has been observed (18). Furthermore, no comment on severity as a separate entity from potency was given in studies aimed at developing a dose–response variant of the GPMT to assess the strength of a sensitizer (19). Indeed, it is the general experience of the authors that, in guinea pig tests, for any given skin sensitizer, the higher the induction dose, the lower the elicitation threshold (data not shown).

The other primary test now used for the predictive (hazard) identification of skin-sensitizing chemicals is the local lymph node assay (LLNA) (20, 21). In a further development, the LLNA also proved to have the ability to
measure the relative skin sensitizing potency of contact allergens (8, 9, 22). However, there is no readout from the LLNA that provides evidence on severity and that is distinct from the potency of a sensitizer. This observation can also be extended to respiratory sensitizers, which are also generally positive in this assay (2).

The clinical expression of severity in humans

In terms of clinical severity, allergic contact dermatitis can range from mild erythema and/or slight sensory effects, through itching, profound erythema, perhaps with oedema, to oozing, blistering skin eruptions. Regarding respiratory allergy, symptoms can progress from mild rhinitis to frank and severe asthma. The dilemma for regulatory toxicology, and the question addressed here, is whether differences in the severity of allergic reactions are reflective of anything more than the extent of exposure to the inducing allergen at the induction and/or elicitation phases.

Inevitably, the evidence that exists to inform the above question derives largely from the study of allergic contact dermatitis in humans. A substantial body of data is available from studies in which skin sensitizers have been used to investigate parameters associated with both the induction and elicitation of allergic contact dermatitis in humans (23–27). Within this body of work, however, there is only limited information on the relationship between induction and elicitation thresholds. What is not available from these studies is any information on the intensity of individual elicitation reactions.

In reality, the only information that exists on the severity of individual responses to skin sensitizers in humans derives from diagnostic patch test results. Patients with a suspected diagnosis of allergic contact dermatitis undergo patch testing (28). The resultant reactions are scored on a simple scale, which is barely more detailed than that for the guinea pig tests mentioned above, and is shown in Table 1 (29). Unfortunately, although the reaction intensity may, in part, reflect the degree to which sensitization has been induced in the subject, the inducing stimuli are often very poorly characterized. Also, the severity of the presenting eczema may depend greatly on the presence of concomitant exposure to irritants, which are well known to enhance the expression of allergy (30, 31). Nevertheless, there is an interesting example in the form of the known contact allergen \textit{p}-phenylenediamine. This chemical has a range of uses, including as an active ingredient in hair dyes and in so-called ‘temporary henna’ tattoos. This latter usage has a tendency to induce a highly sensitized state, and to produce, as a consequence, very severe reactions upon application of a diagnostic patch test, such that it has been proposed that a 100-fold lower test concentration is more appropriate for individuals who have had a ‘temporary henna’ tattoo (32). As \textit{p}-phenylenediamine is acknowledged be an extreme skin sensitizer, this situation is not unexpected (33–35). Despite this, at low levels of sensitization/low elicitation concentrations, skin responses to \textit{p}-phenylenediamine are weak, or even undetectable, so that, for example, those with a low degree of induction (the + reactors on the scale shown in Table 1) may be able to continue to use hair dye without significant problems (36) – although this does not mean that their degree of sensitization will not increase with the continuing exposure. In contrast, none of the more highly sensitized individuals (the +++ reactors in Table 1) continue to dye their hair, because of the strength of the allergic reaction that they experience. Very occasionally, \textit{p}-phenylenediamine has even been reported to produce anaphylactic-type reactions, albeit extremely rarely (reviewed in 37). The interesting question is whether skin sensitizing chemicals of lower potency can also be associated with the elicitation of severe reactions.

\textit{Methyldibromo glutaronitrile} is an order of magnitude less potent than \textit{p}-phenylenediamine, according to predictions from the LLNA (38). However, it has been associated with severe reactions in many patients (39). \textit{Methyldibromo glutaronitrile}, of course, is still regarded as a strong sensitizer, but many substances possess a very much lower potential. For example, in the LLNA, it is reported that skin sensitizers span a potency range of at least five orders of magnitude (8, 9, 22, 38). A rather weaker allergen, at least on the basis of the LLNA, is \textit{hydroxyisohexyl 3-cyclohexene carboxaldehyde}, which has produced a surprising amount of skin allergy (40, 41). Individually, these reactions have ranged from mild to severe.

Finally, it is worth noting that a number of extremely weakly sensitizing substances (i.e. too weak to be classified under GHS-associated regulations), including

---

Table 1. Allergic contact dermatitis patch test scoring scheme

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No reaction</td>
</tr>
<tr>
<td>1</td>
<td>Doubtful reaction, may be irritant – faint erythema only</td>
</tr>
<tr>
<td>2</td>
<td>Weak allergic reaction – erythema, infiltration, possibly papules</td>
</tr>
<tr>
<td>3</td>
<td>Moderate allergic reaction – erythema, infiltration, papules, vesicles</td>
</tr>
<tr>
<td>4</td>
<td>Strong allergic reaction – intense erythema and infiltration, coalescing vesicles</td>
</tr>
</tbody>
</table>

© 2012 John Wiley & Sons A/S

Contact Dermatitis, 67, 3–8
the parabens family of preservatives and long-chain fatty alcohols [such as cetearyl (cetostearyl) alcohol], have been associated with chronic and often severe skin reactions. However, this is undoubtedly because of the use of these materials as topical medicaments that are applied to stasis ulcers (reviewed in 42).

The clinical diagnosis of allergic contact dermatitis depends heavily on the process of diagnostic patch testing, in which the most common and suspected causes of eczema in a patient are applied to the skin of the back for 48 hr (28). The concentration of each chemical normally approaches the maximum non-irritant concentration, but this also has to be adjusted to allow for the potency of the agent, so that active skin sensitization can be avoided. Thus, the use of a low concentration (e.g. 0.01% for methylchloroisothiazolinone/methylisothiazolinone), rather than a very high concentration (e.g. 20% for neomycin), is consistent with a greater relative sensitizing potency for the former. However, clinical experience indicates that both can produce mild to severe eczema. Furthermore, it has recently been recognized that, in fact, contact allergens of varying potency have very similar dose–response profiles for the elicitation of allergic contact dermatitis (43).

With respect to respiratory allergy to chemicals, it is really not possible to say with certainty whether relative sensitizing potency plays any role in the severity of the reactions observed. However, by analogy with allergic contact dermatitis, the expectation is that a major driver is the extent of exposure. Again, the implication is that assessment of the severity of allergic reactions cannot be used with confidence as an indicator of sensitizing potency.

**Conclusions**

The GHS produced under the auspices of the UN took into account many differing regulatory requirements from around world. These included regulations that had long been out of date. Probably as a consequence of ‘political bargaining’, certain anachronistic elements found their way into the documents. These included the concept of severity in sensitization. Whereas it is possible, at least for skin sensitization, to make an estimation of the relative potency of a sensitizer (38, 44), for both skin and respiratory sensitization there is no evidence that some sensitizers produce more severe reactions than others. In reality, the evidence points overwhelmingly to the severity of an expressed allergic reaction being a function of the dose to which the individual has been exposed and the extent to which sensitization has already been induced in that individual. As a consequence, it is hard not to conclude that the concept of severity in regulatory toxicology for sensitization is entirely inappropriate, a point that could at least be addressed in future updates of the regulations (7) and related guidance material (45). Severity is, in reality, a function of intrinsic potency and dose. Thus, although severity remains within the GHS wording, we will need to remain vigilant to ensure that the concept is not misused, so leading to inappropriate classification and labelling decisions.

**References**

16 Magnusson B, Kligman A M. Allergic Contact Dermatitis in the Guinea Pig. Identification of Contact Allergens, Copenhagen, Munksgaard, 1981.
30 Allenby C F, Basketter D A. A model to simulate the effect of detergent on skin and evaluate any resulting effect on contact allergic reactions. Contact Dermatitis 1990: 23: 291.