Available online at www.sciencedirect.com
ScienceDirect
Food and Chemical Toxicology 45 (2007) 691–701
www.elsevier.com/locate/foodchemtox

Hazard characterisation in food allergen risk assessment: The application of statistical approaches and the use of clinical data

R.W.R. Crevel [a,*], D. Briggs [a], S.L. Hefle [b], A.C. Knulst [c], S.L. Taylor [b]

[a] *Safety and Environmental Assurance Centre, Unilever, Bedford MK44 1LQ, UK*
[b] *Department of Food Science and Technology and Food Allergy Research and Resource Program, University of Nebraska, Lincoln, USA*
[c] *Department Dermatology/Allergology, University Medical Centre, Utrecht, The Netherlands*

Received 19 July 2005; accepted 12 September 2006

Abstract

A structured approach to assess the risk to allergic individuals from food allergens requires as a first step the experimental measurement of minimum eliciting doses in a population that is as representative as possible of the relevant allergic population, using a standardised protocol. These doses are established in controlled challenge studies, but logistical and statistical constraints mean that a proportion of the allergic population may still be at risk of reacting at doses below those which have been or could feasibly be tested. However, statistical modelling of the dose distribution resulting from such challenges permits inferences to be drawn about the proportion of allergic individuals that are likely to react to specified (low) amounts of residual allergen in food. However, different statistical models, which all provide good fits to the experimental data yield different values outside the experimental range. Consequently, the outputs from these models require a form of validation, which demonstrates how close the predictions are to reality. In addition to characterisation of the hazard, for each allergenic food this validation requires information about exposure to undeclared allergen, the actual number of reactions taking place in the wider allergic population, and the prevalence of allergy to that food.

Keywords: Minimum eliciting dose; Modelling

1. Introduction

Food allergy is defined as an adverse reaction to food, mediated by the immune system. Of greatest concern is IgE-mediated food allergy, in which the sufferer produces IgE antibodies to one or more proteins present in the food. Subsequent encounters with the food can provoke reactions which range in severity from tingling in the mouth to death through circulatory collapse, often aggravated by bronchoconstriction (anaphylactic shock). Recent data from the USA (Sampson, 2004) and Europe (Kanny et al., 2001) indicate that around 3–4% of the total population suffer from this type of food allergy, making it a significant public health problem, although the percentage at risk of a severe life-threatening reaction is considerably lower though not easily estimated.

The amount of protein reported to provoke an objective allergic reaction ranges from perhaps a tenth of a milligram up to grams, and sometimes tens of grams; considerable individual variability exists among food-allergic individuals. However, these amounts are not well defined for many allergenic foods. At the level of the allergic population, the lower the dose, the less severe the symptoms, and the lower the proportion of reactive sufferers. Yet considerable

Abbreviations: DBPCFC, Double Blind Placebo Controlled Food Challenge; ED, Eliciting Dose; EFSA, European Food Safety Authority; FAO/WHO, Food and Agriculture Organisation/World Health Organisation; ICMSF, International Commission on Microbiological Specifications for Foods; LOEL, Lowest Observed Effect Level; NOEL, No Observed Effect Level; USDA, United States Department of Agriculture.

* Corresponding author. Present address. ILSI Europe, ILSI Europe A.I.S.B.L., Av. E. Mounier 83 – Box 6, B-1200 Brussels, Belgium. Tel.: +32 (0)2 771 00 14; fax: +32 (0)2 762 00 44.
E-mail address: rene.crevel@unilever.com (R.W.R. Crevel).

0278-6915/$ - see front matter
doi:10.1016/j.fct.2006.09.005

uncertainty exists regarding the lowest amount to which sufferers will react, the proportion of sufferers reacting to a defined dose, as well as the relationship between dose and severity for any given individual. Assessing the public health impact of the small amounts of residual allergens that are not part of the formulation of a manufactured food product but may be present due to the production process, represents therefore a considerable challenge to risk managers in public authorities and industry.

A structured approach to assess the risk from food allergens to individuals who know they are allergic requires as a first step the experimental measurement of the highest dose not eliciting a reaction (the No Observed Effect Level, or NOEL) in a population which is as representative as possible of the relevant allergic population, using a standardised protocol. However, such experimental approaches have several limitations. These include the feasibility of food challenges in adequate numbers of people, the criteria for inclusion in such studies, the extent to which challenges in a clinic can reproduce the response of free-living individuals, and the differences between real foods and the matrices used to mask the allergenic food. Ultimately, these studies do not allow direct estimation of the amounts of allergen that could provoke reactions in the small proportion of the population who react to doses lower than any tested, as discussed later. That section of the allergic population is nevertheless of legitimate concern to risk managers in industry and public authorities, and allergic patients naturally wish for an assurance that they are not being put at risk. To overcome this problem in other areas of food safety (e.g. microbiological contamination), modelling approaches have been used successfully (ICMSF, 2002). In 2002, Bindslev-Jensen et al. proposed a modelling approach (Bindslev-Jensen et al., 2002), and illustrated it for four common allergenic foods, using data published in the peer-reviewed literature. The authors stressed the preliminary nature of their findings, including the numerical values derived. The publication thus marked a significant step forward in raising the issue of the generation and use of clinical data for allergen risk assessment, but many issues remained unresolved, as highlighted by the authors.

The purpose of this paper is to review how statistical approaches can contribute to the assessment and thereby the management of the risk to people already sensitised to specific food allergens from exposure to them. It discusses how such approaches can help to characterise the hazard, and specifically considers how the limitations already identified can be best overcome to provide an approach that can be used with confidence by risk assessors and managers. Based upon a statistical analysis of clinical data, a 3-stage approach is proposed to enable the derivation of tolerable upper levels for allergens.

2. Managing risks from food allergens

Managing allergen risks is a shared responsibility of all the stakeholders and requires a consistent approach across industry in order to be effective. The food industry must ensure that allergenic ingredients are declared and that residual inadvertent allergen levels are safe for the vast majority of allergic consumers. Health professionals diagnose and advise patients, who in turn must exercise due care in their food and product choices, while regulators must balance the interests of all stakeholders and ensure compliance.

Because of the uncertainties associated with food allergy and food allergens, current approaches to protect allergic consumers have focussed on allergenic hazards rather than attempting to assess the risk. Thus legislation in several regions now makes mandatory the declaration of specified allergenic ingredients whenever present in a manufactured food. This approach is generally effective in managing the risk from ingredients present in meaningful quantities, although it is arguably unhelpful for minor ingredients added in minuscule amounts, for instance, because they are minor constitutents in another ingredient. However, the risks associated with allergens do not arise solely from their use as ingredients. Allergenic constituents can be present through cross-contact during manufacturing, or through their presence in raw materials. Thus, allergenic ingredients such as peanuts, milk, eggs and wheat are used extensively in food manufacture, and guaranteeing total absence of all such constituents from a product when not present as ingredients is often not a practical proposition. Managing the risk to which they give rise on the basis of hazard has led to extensive use of precautionary (''may contain'') labelling. Allergic individuals dislike precautionary labelling because they are unsure about the basis for its use and it reduces their food choices considerably (Sicherer et al., 2001; Avery et al., 2003). Arguably it also puts them at increased risk not only directly through devaluation of the label (Mills et al., 2004) but also through adverse effects on nutritional status (Christie et al., 2003). Aside from the impact on the allergic consumer, the hazard-based approach to allergen management can also cause problems for the food industry by driving the implementation of measures which may go beyond what is necessary to make food safe for allergic consumers.

Empirical data suggest (but certainly do not prove) that exposure to very low levels of protein from allergenic foods is unlikely to pose risks to the vast majority of food-allergic consumers. For example, USDA grain standards allow the presence of small amounts of soy in corn and wheat in oats without declaration on product labels. Soy-allergic individuals are not known to react from consuming corn (maize) products although this has never been critically examined by controlled challenge studies. Many other examples exist of low-level exposures of consumers to ingredients derived from allergenic foods that are not typically declared on product labels and complaints from allergic consumers about such products have not been noted; examples include the widespread use of soy lecithin containing traces of soy protein as a release agent in baking operations. A risk-based approach, founded on characterisation of the relationship

between exposure to low levels of residual allergen, and frequency and severity of likely reactions, offers better outcomes for all stakeholders. Food manufacturers need such information in order to optimise their allergen control procedures without jeopardising their economic viability. Defined rules for the application of precautionary labelling should reduce its extent while increasing its credibility, to the benefit of allergic consumers. Similarly such information would facilitate the role of regulatory authorities in balancing the interests of different stakeholders.

3. Risk assessment, thresholds and eliciting doses

The process of risk assessment is conventionally broken down into four separate stages: hazard identification, hazard characterisation, exposure assessment and risk characterisation (FAO/WHO, 1995; FAO/WHO, 1997; Codex Alimentarius Commission, 2003).

This framework can be applied to food allergy. Thus, for food allergens, the hazard is implicitly identified insofar as it is the potential to cause an allergic reaction. However, despite the growing amount of data becoming available, the hazard is poorly characterised, to the extent that amounts of allergen below which no one reacts cannot be defined currently. Thus, based on a review of the existing data, Taylor et al. (2002) concluded that ''thresholds for common allergenic foods are finite, measurable and above zero. However, attempting to reach consensus on the threshold doses for peanut, egg, cows' milk, fish, and mustard on the basis of the existing data would probably be premature''. The European Food Safety Authority (EFSA, 2004) reached a similar conclusion more recently. Furthermore, the relationship between the dose of the allergenic food and the nature and severity of the response is even less well defined, and for ethical reasons, will be even more difficult to establish.

Modelling of the population distribution of minimum eliciting doses, based on clinical food challenge data, is an attempt to overcome this problem by introducing an element of hazard characterisation at the population level. The modelling approach previously described (Bindslev-Jensen et al., 2002) proposed to address these issues, but also highlighted a need to review the terminology to make it reflect more accurately the state of knowledge.

Allergic responses, in common with other adaptive immune responses, consist of two phases: sensitisation and elicitation. Thresholds can apply to both phases. However, nothing is known about thresholds of sensitisation to food proteins in human beings and in practice, the term ''threshold'' is only used in relation to the elicitation phase. Food allergens differ from materials subject to conventional toxicological risk assessments insofar as they do not provoke reactions at any reasonably attainable dose in the vast majority of the population. Many also form a significant part of the diet and contribute significantly to nutrition. No imperative therefore exists to identify a threshold below which sensitisation to common allergenic foods would not occur. This paper will therefore also only address thresholds of elicitation.

The Concise Oxford English Dictionary (9th Edition) defines threshold (Physiology) as ''a limit below which a stimulus causes no reaction''. Kroes et al. (2000) define threshold in toxicology as a dose at, or below which, a response is not seen in an experimental setting. They also highlight that economic and practical limitations make it impossible strictly to derive absolute values for thresholds. The term is also used with other meanings in the same context to signify an exposure so low as to be of no toxicological concern, as discussed by Edler et al. (2002). However, plotting the distribution of patients' minimum eliciting doses, determined in controlled clinical challenge studies does not necessarily yield a threshold below which no member of the wider allergic population would react because of the obvious uncertainty not only in locating the very most sensitive persons for challenge trials, but also in ensuring that the challenged population is representative of the whole allergic population. We therefore propose the term ''eliciting dose'' (ED), where ED_p refers to the dose of allergen that produces a response in p% of the allergic population. This terminology better reflects the fact that realistic clinical challenge studies can only define the proportion of the allergic population likely to react to a specified dose, rather than establish a definitive amount below which no allergic person would react.

Systematic consideration of each element in risk assessment highlights the need to understand exposure in order to fully characterise the risk. This will be discussed further when considering the application of statistical modelling in a later section, although a full discussion of exposure assessment is beyond the scope of this paper.

4. Approaches to hazard characterisation as a starting point for risk management

The previous sections identified the need to establish benchmarks for allergen risk management. Such benchmarks could be chosen in several ways, ranging from a review of the relevant published or unpublished data using expert judgement, to the experimental establishment of minimum eliciting doses in well-controlled challenge studies, accompanied by the use of statistical methodology to make inferences beyond the experimental range.

4.1. Expert review of clinical and non-clinical data

The simplest way to choose a benchmark is probably to review the available literature for the relevant allergenic food, and use the lowest amount reported to have caused an observable (objective) reaction in a reported allergic incident as the benchmark. Expert judgement can be invoked to refine this definition and to allow for uncertainty. While this approach can appeal through its simplicity, it possesses a number of drawbacks. The first is that in many such case reports, the amount of allergenic food ingested is not

usually accurately known. Other foods eaten at the same time complicate the identification of the causative food. Another difficulty is reliance upon case reports when there is little indication about how frequent such events are, and thus whether the risk management actions which this approach entails are commensurate with the benefit that they bring. The individuals suffering those reactions are also often not very well characterised, making it difficult to establish how representative they are of the wider allergic population. Typically, only the most unusual case reports merit publication. The benchmark would therefore need to be revised whenever new findings are reported, irrespective of their frequency, which could pose problems for consistency among risk managers. The chosen values are thus to some extent arbitrary and it is difficult to demonstrate how they are derived from the data, and therefore it is also difficult to justify their use as a basis for action. Even where quantitative information is reported, its value may be compromised by the limitations of analytical methods.

An alternative approach could be to use the NOELs or LOELs obtained from controlled clinical challenge studies, and apply an uncertainty factor to them in order to derive a value for use in risk management of the allergen of interest, as is the practice in other areas of toxicology. One advantage is the simplicity of this approach, but it uses only a single data point from the challenge studies, and it suffers from the same problems as other methods with regard to selection of appropriate uncertainty factors. The extent to which it protects the allergic population is also difficult to establish, compared to a statistical approach because it does not allow statements to be made about likely reaction rates to doses outside the experimental range. Even if published results from controlled, low-dose challenge trials are used to gather data for establishment of minimum eliciting doses, difficulties exist in assessing the exposure dosage because of the variability in forms of the food that are used e.g. non-fat dry milk, fluid milk, infant formula (Taylor et al., 2002; Grimshaw et al., 2003).

4.2. Application of the Binomial theorem

A more systematic approach is to use data from double-blind placebo controlled food challenges (DBPCFC) to define the proportion of allergic individuals who will react to a specified dose. Using a Binomial probability approach and the concept of statistical confidence intervals, this method can provide a degree of assurance that less than a specified proportion of the allergic population will react to each dose tested. This same logic can be applied to the NOEL, in providing assurance that less than a specified proportion of the allergic population will react to the NOEL.

For the NOEL in a study including 29 subjects, statistically it can be said that there is 95% confidence that less than 10% of the allergic population will react. Food challenge studies based on these numbers have become a *de facto* standard, by analogy with the unofficial US standard for the acceptance of hypoallergenic infant formulae (Kleinman et al., 1991). They also represent a pragmatic balance between the desired degree of confidence and practicability.

Table 1
Number of participants required to conclude that the reaction rate in the allergic population is significantly less than a specified value (at the 95% level of statistical significance)

Reaction rate (upper confidence limit 95% significance, 1-sided)	Number of participants required	
	…With no reactions	…With 1 reaction
10%	29	46
5%	59	93
1%	299	473

Table 1 shows that the lower the reaction rate about which one wants to make statements, the greater the number of study participants required, such that the study rapidly becomes impracticable. This situation is exacerbated if any reactions at all occur at the lowest dose tested. Therefore in order to make precise statements about low reaction rates in the allergic population, a large number of participants is required with no (or very few) allergic reactions seen at the lowest dose tested.

This approach also only examines the NOEL independently of all other doses, therefore it does not take into account the proportions of reactions at different dose levels (i.e. the shape of the dose distribution curve). It is also only possible to make inferences about the tested doses, while in practice it is desirable to make inferences about lower doses than those tested.

4.3. Statistical distributions

A third way of establishing a benchmark for risk management is to examine the distribution of clinical minimum eliciting doses. This approach permits (in principle at least) inferences to be made about reaction rates to dose levels outside, and especially below, the experimental range. It can thus form the basis for stakeholders to agree a benchmark for risk management through analysis of the frequency-dose distribution at low doses.

A distribution model should provide a good fit for the observed data and be scientifically plausible. Where several models provide good fits, the choice of distribution model is important as the estimated proportion of the allergic population reacting to low dose levels will depend upon the distribution used, as well as on the observed clinical data. Thus, consideration should be given to whether a distribution includes a threshold of no response as a characteristic (i.e. a dose level below which no one will react). If thresholds of no response actually do exist for allergic reactions (albeit at very low dose levels) then a distribution which implies a true threshold might be appropriate (example shown in Fig. 1b). However, data exist from controlled clinical challenge trials with particular foods to indicate that some allergic individuals may react (though not severely) to extremely small quantities of allergenic food, of the order of 0.1 mg and very rarely lower. Clinical case

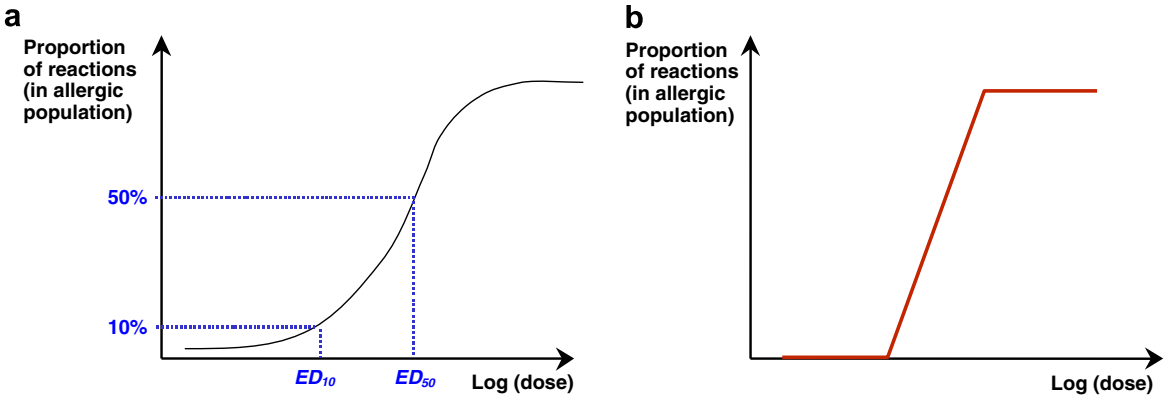

Fig. 1. Graphs showing the difference between an unbounded distribution (no threshold implied) and a bounded distribution, where a threshold of no response is implied.

reports suggest on a much more anecdotal basis that exposure to even lower amounts of an allergenic food can sometimes elicit allergic reactions although the dose is rarely known with any degree of certainty and the cause-and-effect relationship is often speculative. Therefore, while thresholds of no response almost certainly exist, they are predictably difficult to define with confidence unless large numbers of subjects are challenged. The anecdotal evidence while weak thus seems to argue against the existence of a true experimentally measurable threshold, giving support instead for an unbounded distribution (where the curve never intersects the horizontal axis, Fig. 1a). Thus, a no-threshold model (Fig. 1a) is probably a more appropriate description for the distribution of minimum eliciting doses from controlled clinical challenge trials. This is also illustrated by experimental findings such as those obtained with roasted peanut by Wensing et al. (2002) (Fig. 2). The authors used subjective symptoms as challenge endpoints, which is the subject of debate from a clinical perspective, but does not affect this illustration. Thus, the sensitivity analysis presented in the next section of this paper will focus on unbounded distributions.

The ultimate test of any predictive approach such as this one is whether the results of extrapolation outside the experimental zone agree with epidemiological observations on the prevalence and incidence of food allergies. However the data required to validate such findings are generally not available, as discussed later. It is therefore difficult to assess the distributions for their relevance to the wider allergic population, and the impact that the choice of distribution has on predictions needs to be assessed using sensitivity analysis. A no-threshold model implies that some individuals could react at any dose and that 100% protection of the entire allergic population is impossible. This approach therefore raises risk communication issues, but these are beyond the scope of this manuscript.

4.3.1. Sensitivity analysis

Sensitivity analysis provides a way of assessing the impact that the choice of distribution has on predictions,

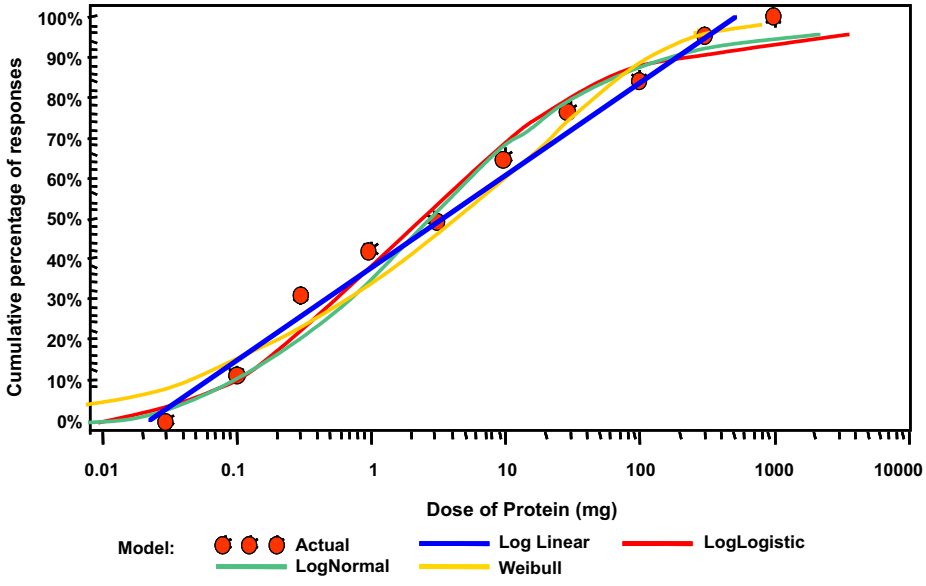

Fig. 2. Representation of challenge data from Wensing et al. (2002), illustrating different statistical fits.

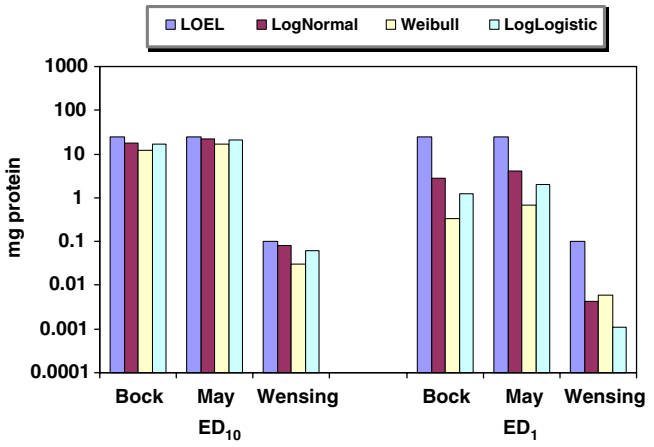

Fig. 3. Bar chart showing the differences in ED_{10} and ED_1 values across distributions, for three peanut studies.

and is an essential component of investigating the robustness of the approach. The equations for the distributions listed here can be found in many statistical textbooks (Johnson et al., 1994).

The distributions examined in this section are the Log-Normal, Weibull and the Log-Logistic. To illustrate, Fig. 3 below show results from these three different distributions fitted to data from three clinical studies on peanut, two of which were examined by Bindslev-Jensen et al. (2002) (Bock et al., 1978; May, 1976). The Lowest Observed Effect Levels (LOELs) are also presented for comparison.

Fig. 3 shows that the differences in ED_{10} value (at the lower boundary of the experimental zone) between the different studies are greater than the differences due to the choice of distribution model. They are thus effectively determined by the data, including the use of subjective endpoints in the Wensing et al. (2002) studies in contrast to the objective endpoints of the earlier studies, rather than the statistical model. Possible reasons for these differences between studies will be discussed in more detail in Section 4.4.

The ED_1 values however (outside the experimental range), are strongly influenced by the choice of distribution model as well as by the data. Estimates of ED_1 vary by an order of magnitude depending upon the choice of distribution model. This parallels the experience with low dose extrapolation models applied in other areas of toxicology, such as cancer models (Edler et al., 2002). This illustrates the importance of choosing a distribution model that both fits the data within the experimental range, as well as one which is scientifically plausible for extrapolation to low doses.

However, the differences caused by choice of distribution model may be relatively small in the wider context of risk assessment. Uncertainty about the prevalence of food allergy, the reactivity of the study population, whether analytical methods have been used appropriately, and food consumption estimates, may all contribute more to uncertainty in risk assessment than the choice of the statistical distribution model. This point remains to be addressed, but is outside the scope of this paper.

4.3.2. Combining data from different studies

Data from clinical studies can in principle be combined, either by merging studies which examine the same allergenic food, or by merging studies which examine different allergenic foods. Combining data from studies which examine different allergenic foods is not appropriate for risk assessment as there is evidence that different allergenic foods and their constituent proteins have different potencies, based on the doses required to elicit responses (European Food Safety Authority, 2004) and therefore would also have different distributions of minimum eliciting doses. For risk assessment, studies should be used which challenge patients with food specific to the particular risk assessment scenario.

Merging studies which examined the same allergenic food, however, has the advantage of combining a wider range of study participants of potentially varying sensitivities, possibly helping to overcome the criticism that the study participants are not representative of the wider allergic population. In addition, the merging of different studies, which may have used different challenge matrices, different dosing regimes, etc might have the advantage of "averaging out" these differences, so that the final merged dataset better resembles the exposure scenarios which the wider allergic population experiences, while increasing the statistical power of the predicted estimates. However, merging studies may create other issues such as the interpretability of the predictions and the influence of differences between experimental protocols. Some of the challenges associated with merging data from different studies have been previously reviewed (Taylor et al., 2002).

Certainly, studies which contain subjects of a similar range of sensitivity, with similar dosing regimes, using similar challenge matrices, etc, could be combined to provide datasets which are based on a greater number of people, and are therefore likely to provide more reliable and repeatable estimates of the relationship between the dose of allergenic food and the rate of allergic reactions. The use of standardised protocols such as that of Taylor et al. (2004) for conducting DBPCFC studies will help in this regard, but testing all forms of an allergenic food may be a practical impossibility.

When conducting a risk assessment however, the most appropriate study/studies should be used, as defined by factors such as the study population and the form of the allergenic food. For example, if a risk assessment is aimed at protecting children, then data from a clinical study that examines children should be used, if available. In addition, the challenge material used in the clinical study should as closely as possible resemble the form of the food that is the focus for the risk assessment. For example, if roasted peanut is the allergen of concern, then data from a clinical study should be used which challenged patients with roasted peanut.

In addition to evaluating appropriate existing studies aimed at determining minimum eliciting doses, the possibility exists to collect data from diagnostic challenges conducted at allergy clinics on individual patients. Although the aim of such challenges is the diagnosis of a specific food allergy, each clinic likely uses consistent food challenge procedures. In fact published studies on minimum eliciting doses may be skewed toward the more highly sensitive patients as opposed to patients who react to more typical diagnostic challenge doses (Sicherer et al., 2000). Inclusion of these additional challenge results might therefore help to make the distribution more representative.

4.4. Application of statistical distributions in food allergen risk assessment

The essence of the approach described in this paper is to use the dose distribution of minimum eliciting doses in a sample of allergic people to make predictions about the likely incidence of reactions to amounts of allergenic food outside the range that can be experimentally tested. Interpretation of the resulting predictions requires consideration of factors specific to the modelling as well as to the experimental design of the studies. The approach relies on good quality clinical data from well-controlled challenge studies. However, even where studies meet the relevant quality criteria many variables can affect the observed distribution of minimum eliciting doses.

Uncertainty in eliciting dose estimates. The statistical approach as detailed here provides an estimate of the dose that elicits a reaction in a specified proportion of the allergic population. However, the clinical data are subject to uncertainty which needs to be accounted for when estimating eliciting doses. This variability in the data can be captured by citing of a confidence interval around an eliciting dose. For allergen risk assessment and management, it is recommended that the lower confidence interval on the dose should be taken as a conservative estimate. This approach is similar to the application of the Benchmark Dose approach as developed by Crump (1984).

Study population characteristics. In the statistical modelling approach, challenge data from a variety of studies are used, subject to criteria such as number of subjects in the study, protocol design and the use of an adequate number of doses. However, the explicit purpose of the model is to extrapolate from the tested population to the allergic population as a whole in order to predict the numbers of individuals likely to be affected by amounts of allergen below those which can be tested experimentally. The implicit assumption behind the model is therefore that the study populations can be considered representative samples of the overall allergic population. Before predictions from the model are used operationally, this assumption needs to be rigorously tested. For instance, it has been taken for granted that individuals who have suffered a severe reaction are almost always excluded from participation in challenge studies on the grounds of safety. Furthermore, severe reactors have usually been equated with those responding to very low doses, i.e. sensitive ones. Both these points must be questioned in the light of recent clinical practice. For instance, by starting at extremely low doses, Wensing et al. (2002) were able to include very sensitive individuals in challenge studies designed to investigate population dose–responses with the most potent allergens such as roasted peanuts. Other clinicians have presented data clearly demonstrating that low dose reactivity is not synonymous with the ability to experience a severe reaction at a low dose (Bindslev-Jensen data, Hefle et al., 2003). Of even greater importance is the method of selection of participants in challenge studies. In almost all cases, they are recruited from the population attending specialist allergy clinics. In many countries, such patients represent only the individuals most severely affected by their allergy, who have been referred by general physicians for specialised care because of the difficulties in managing their condition. In this context, it may be interesting to note the minimum eliciting doses observed in challenges associated with epidemiological studies (e.g. Zuberbier et al., 2004), compared to those seen in studies based on clinic populations. Subject to the limitations associated with comparing different populations, the findings support the argument that the clinic population is indeed a more sensitive and reactive sub sample of the whole allergic population. This issue might be answered by comparing the frequency of severe and mild symptoms reported in the test population and in the allergic population at large. This in itself makes certain assumptions, of course, specifically that the spectrum of exposure scenarios does not differ between the two groups, and neither does the allergen exposure that provoked their reactions.

The mathematical considerations in extrapolating beyond the experimental zone have already been discussed. However extrapolation from the clinical environment to the community is also a key consideration, with environmental factors such as exercise, alcohol consumption, and food processing are all known to modulate the way that allergic individuals react to contact with allergen (Aihara et al., 2002; Kopper et al., 2005). Since allergic individuals almost always suffer their reactions in the community and not in a clinic setting, an understanding of the relationship between clinical predictions and reactions that are seen in the community is essential in order to validate this approach.

Interpretation of an adverse reaction. Reactions to an allergenic food may be subjective, i.e. reportable only by the subject being challenged, or objective, and therefore capable of independent verification by an outside observer. In challenge studies, subjective reactions usually occur before objective, independently observable ones, and indeed, in the community, such reactions are often treated by allergic people as warning signs (Sicherer et al., 2000; Wensing et al., 2002). In some challenge studies, subjective symptoms are the only ones reported for some subjects. It is debatable whether eliciting doses based on subjective symptoms can or should be grouped with eliciting doses

based on objective symptoms for the purpose of modelling the responses.

Interpretation of the dose data. Individual eliciting doses have been defined in different ways in different challenge studies. For instance, Wensing et al. (2002) cite the lowest individual dose causing a reaction as the eliciting dose whereas challenge studies conducted by the Food Allergy Research and Resource Program (FARRP) of the University of Nebraska (Taylor et al., 2004) cite the cumulative dose up to the point where the subject reacted. Using such different dose metrics will affect the shape of the distribution of minimum eliciting doses. The latter (cumulative dose) approach implicitly assumes that all exposures taking place within a certain time frame are summed, whereas the former approach treats individual exposures as independent events without influence on each other. The validity of one approach or the other from a biological point of view depends on factors such as the time spacing of doses, and the matrix in which the allergen is masked. A consensus conference determined that data from such trials should be presented both as discrete and cumulative doses because considerable debate could ensue about the most appropriate approach (Taylor et al., 2004). Currently no data exist to support one approach or the other. Nevertheless, considering each subsequent dose as an individual dose represents a conservative approach to risk assessment, since it leads to a lower minimum eliciting dose.

Nature of the challenge material. Both the food and the way in which it is presented may also affect the doses at which reactions are observed. For instance, proteins from roasted peanut bind with greater affinity to specific IgE than their counterparts from raw peanuts (Chung et al., 2003; Maleki and Hurlburt, 2004). Minimum eliciting doses appear to be correspondingly lower, as illustrated by the contrasting results of the studies of Bock et al. (1978) (raw peanut) and Wensing et al. (2002) (roasted peanuts), although comparison of these two studies is complicated by the choice of different challenge endpoints, as discussed earlier. Different cultivars of the same plant can be another source of variability. For instance, the amount of the allergen Mal d1 varies considerably between apple varieties (Vieths et al., 1993, 1994, 1996). The food used for challenge may not be the only source of variability. Recently Grimshaw et al. (2003) showed that the fat content of the matrix used to disguise peanut flour altered considerably the minimum eliciting dose in several peanut-allergic individuals. Their data suggest that these changes occurred because the allergen appeared to be released more slowly from a higher fat matrix. Another aspect of the matrix, which has not received as much attention as it deserves, is the effectiveness of the blinding. Given that allergic responses can occur as conditioned responses, it is essential to verify that the matrix truly disguises the challenge food in every respect (flavour, smell, texture, etc). This can only usually be achieved with the use of a sensory panel using standard procedures (Huijbers et al., 1994; Ronteltap et al., 2004; Vlieg-Boerstra et al., 2004).

Many of the issues outlined in this section are however not unique to allergen hazard assessment, with established hazard assessment approaches in other areas of toxicology suffering from similar knowledge gaps. In fact a major advantage of allergen hazard assessment is the availability of clinical data rather than only data from animal or in-vitro models, removing the need for inter-species extrapolation. The authors believe that the modelling described provides an alternative and complementary approach to consider in gaining an understanding of the relationship between a dose of allergen ingested and the likely reaction rate in the allergic population. However, this hazard assessment remains a first step in allergen risk assessment and must be informed by a wider knowledge of food allergy and its impact. Full risk assessment and management also requires implementation of the other stages of risk assessment, including a thorough assessment of exposure. An important benefit of using a statistical modelling approach in allergen risk assessment lies in the transparency of the conclusions drawn from the clinical data, enabling a more productive interaction between different groups of stakeholders.

5. Conclusions and recommendations

Allergen risk management must balance the needs of several different groups with potentially divergent interests. The most sensitive allergic individuals ideally need absolute assurance of an absence of potentially provoking amounts of the relevant allergenic protein in products, at the same time as needing a wide choice of products that they can consider safe. This could in theory be achieved by the complete elimination of allergenic ingredients from food manufacturing, or specific manufacturing facilities. However, most allergenic ingredients are a good source of nutrients and form a valuable part of the diet for the general population. Furthermore, there is insufficient evidence about the relationship between exposure to allergenic ingredients and the prevalence of allergy to know whether substitution with other ingredients would achieve the public health aim of reducing the incidence of food allergy. In fact, introduction of certain novel protein sources in certain locales is associated with an increase in allergic sensitisation to the novel food; the best example may be lupine flour in the EU introduced in part to replace genetically modified soybeans (Fernandez et al., 2002). Development of allergen management policies and procedures by food manufacturers contributes significantly to protection of allergic consumers by ensuring that the inadvertent presence of allergenic ingredients in products is minimised (Taylor and Hefle, 1997; Huggett and Hischenhuber, 1998; Crevel, 2002). However, the nature and complexity of food manufacturing operations imposes limits on such approaches which can only be overcome at significant cost, which is ultimately borne by all consumers. These considerations have led significant numbers in the food industry towards the concept, and in many cases the use, of defensive (precau-

tionary) labelling. This type of labelling, while initially welcomed by allergic consumers, has been devalued through perceived over-use and inconsistent application. Indeed it now presents the allergic consumer with a dilemma: observe the labelling and face much decreased choice or ignore it and risk a reaction to the food.

The approach we propose offers an alternative whereby the implications of specified amounts of allergenic ingredient in a product can be defined and used as a basis for risk assessment, subsequently aiding risk management and communication. It recognises that for allergens a NOEL experimentally determined in challenge studies probably still leaves too high a level of uncertainty for risk managers, both in industry and in public health. The approach takes the following steps:

1. Good quality data from well-controlled challenge studies constitute the foundation of the approach. Such studies should be conducted using an agreed low dose protocol such as that proposed by Taylor et al. (2004).
2. A tolerable limit for the proportion (p) of allergic individuals who might react is established for the allergens of interest. The proportion p is defined in consultation with other stakeholders (regulators, patients, industry, clinicians). This part of the process requires other factors to be accounted for such as the prevalence of allergy to the allergen of interest, the actual rate of reported reactions, and their severity.
3. Data from a challenge study are analysed and a statistical distribution is fitted to the study data. The model is used to predict the ED_p and the lower confidence interval on the ED_p for the p defined in step 2.

Precedents from other areas of food safety support this approach, although there are distinct differences, related to the type of data that are available or can be generated in each case. Thus Buchanan et al. (1997) estimated a conservative dose–response relationship for listeriosis using surveillance data on the number of cases, together with food sampling data on *Listeria* contamination of a food responsible for most such cases. In our approach, we use statistical distributions to describe data from clinical challenge studies performed under carefully controlled conditions. In microbiological terms, such data are more akin to controlled feeding studies with defined numbers of microorganisms in human volunteers. As argued earlier, the process of selecting for challenge studies usually results in a bias towards the more severely affected of the allergic population. This arises because volunteers are chosen from the population which attends tertiary referral clinics, and who are therefore sufficiently motivated by the impact of their condition to visit a clinical expert. Limitations exist, of course, when using such data as a basis for risk management, because it does not take into account some of the factors which modulate the allergic response, such as alcohol intake, exercise, etc. However, it can be argued that compared to other areas of toxicology, these data present the major advantage of being generated in the species of interest. Currently available data do not permit a direct application of the methods proposed in Buchanan et al. (1997). Specifically, data on the incidence of even severe allergic reactions to foods are not reliable, and usable published data on the distribution of undeclared allergen in food are very limited. Nevertheless, the next step in the development of the proposed approach is to validate the predictions by comparing the number of predicted reactions with the number actually observed in the community.

We believe that this approach offers an explicit way of assessing the hazard from food allergens for both the industry and the regulatory community, together with the prospect of clearer communication with other stakeholders, particularly allergic consumers. Particularly when further refined, it will provide information which can be used to examine different allergen control scenarios, in terms of the probability of provoking reactions attached to defined levels of cross-contact allergen. It could also, for instance, contribute to the setting of food safety objectives for specific food allergens, as has been done for microbiological hazards (ICMSF, 2002). This would help provide a sound basis for application of the allergen labelling legislation which has recently been implemented in both the European Union (Directive 2003/89/EC) and the United States (Food Allergen Labeling and Consumer Protection Act of 2004).

Acknowledgements

A contribution of the University of Nebraska Agricultural Cooperative Extension, Lincoln, Nebraska, USA 68583. Extension Journal Series No. 1029.

The authors wish to thank Joanne Dick (Unilever R&D) for invaluable statistical assistance, and the ILSI Europe Eliciting Doses Expert Group for helpful discussions.

This work was supported by a grant from the Food Allergy Task Force European branch of the International Life Sciences Institute (ILSI Europe). Industry members of this task force are Ajinomoto Europe SAS, Amylum Group, Barilla G&R F.Lli Spa, Bayer CropScience, Coca-Cola Europe, Eurasia and Middle East, Danisco Cultor Innovation, Group Danone, H.J. Heinz, Kraft Foods, MasterFoods, Monsanto UK Ltd., Nestlé, Numico Research, Pepsico International, Unilever Research, Wild Flavour/Ingredients. For further information about ILSI Europe, call +32 2 771 00 14 or email: info@ilsieurope.be. The opinions expressed herein are those of the authors and do not necessarily represent the views of ILSI and ILSI Europe.

References

Aihara, M., Miyazawa, M., Osuna, H., Tsubaki, K., Ikebe, T., Aihara, Y., Ikezawa, Z., 2002. Food-dependent exercise-induced anaphylaxis: influence of concurrent aspirin administration on skin testing and provocation. Br. J. Dermatol. 146 (3), 466–472.

Avery, N.J., King, R.M., Knight, S., Hourihane, J.O., 2003. Assessment of quality of life in children with peanut allergy. Pediatr. Allergy Immunol. 14, 378–382.

Bindslev-Jensen, C., Briggs, D., Osterballe, M., 2002. Can we determine a threshold level for allergenic foods by statistical analysis of published data in the literature? Allergy 57, 741–746.

Bock, S.A., Lee, W.Y., Remigio, L.K., May, C.D., 1978. Studies of hypersensitivity reactions to foods in infants and children. J. Allergy Clin. Immunol. 62, 1327–1334.

Buchanan, R.L., Damert, W.G., Whiting, R.C., Schothorst, M.V., 1997. Use of epidemiologic and food survey data to estimate a purposefully conservative dose–response relationship for *Listeria monocytogenes* levels and incidence of listeriosis. J Food Prot 60 (8), 918–922.

Christie, L., Hine, J., Parker, J.G., Burks, A.W., 2003. Food allergies in children affect nutrient intake and growth. J. Amer. Diet Assoc. 102 (11), 1648–1651.

Chung, S.Y., Butts, C.L., Maleki, S.J., Champagne, E.T., 2003. Linking peanut allergenicity to the processes of maturation, curing, and roasting. J. Agric. Food Chem. 51 (15), 4273–4277.

Codex Alimentarius Commission, 2003. Draft Principles for the Risk Analysis of Foods Derived from Modern Biotechnology, at Step 8 of the Elaboration Procedure. ALINORM 03/34, Appendix II. Codex Alimentarius Commission, Joint FAO/WHO Food Standards Programme, Food and Agriculture Organisation, Rome. Available from ftp://ftp.fao.org/codex/alinorm01/al0134ae.pdf.

Crevel, R.W.R., 2002. Industrial dimensions of food allergy. Biochem. Soc. Trans. 30 (Pt 6), 941–944.

Crump, K.S., 1984. An improved procedure for low-dose carcinogenic risk assessment from animal data. J. Environ. Pathol. Toxicol. Oncol. 5, 339–348.

Edler, L., Poirier, K., Dourson, M., Kleiner, J., Mileson, B., Nordmann, H., Renwick, A., Slob, W., Walton, K., Würtzen, G., 2002. Mathematical modelling and quantitative methods. Food Chem. Toxicol. 40, 283–326.

European Food Safety Authority, 2004. EFSA Journal 32, 1–197, Opinion of the Scientific Panel on Dietitic Products, Nutrition and Allergies on a request from the Commission relating to the evaluation of allergenic foods for labelling purposes.

FAO/WHO, 1995. Application of Risk Analysis to Food Standards Issues, Report of the Joint FAO/WHO Expert Consultation, Geneva, Switzerland, 13–17 March 1995, WHO/FNU/FOS/95.3. World Health Organisation, Geneva. http://www.who.int/fsf/Documents/applira.pdf.

FAO/WHO, 1997. Risk Management and Food Safety, Report of a Joint FAO/WHO Consultation, Rome, Italy, 27 to 31 January 1997, FAO Food and Nutrition Paper 65. Food and Agriculture Organisation of the United Nations, Rome. Available from: <http://www.fao.org/docrep/W4982E/W4982E00.htm>.

Fernandez, C., Rodriguez, J., Crespo, J.F., 2002. Hypersensitivity to lupine. Food Allergy Intoler. 3, 168–176.

Grimshaw, K.E., King, R.M., Nordlee, J.A., Hefle, S.L., Warner, J.O., Hourihane, J.O., 2003. Presentation of allergen in different food preparations affects the nature of the allergic reaction – a case series. Clin. Exp. Allergy 33 (11), 1581–1585.

Hefle, S.L., Christie, L., Sicherer, S.H., Althage, K.A., Burks, A.W., Sampson, H.A., Mofidi, S., Noone, S., Michaelis, L., Strobel, S., Hourihane, J.O'.B., Nordlee, J.A., Taylor, S.L., 2003. Threshold dose for egg allergy determined by oral challenge. In: Conference Proceedings IFT Meeting 2003.

Huggett, A.C., Hischenhuber, C., 1998. Food manufacturing initiatives to protect the allergic consumer. Allergy Suppl 53 (46), 89–92.

Huijbers, G.B., Colen, A.A., Jansen, J.J., Kardinaal, A.F., Vlieg-Boerstra, B.J., Martens, B.P., 1994. Masking foods for food challenge: practical aspects of masking foods for a double-blind, placebo-controlled food challenge. J. Am. Diet Assoc. 94, 645–649.

International Commission on Microbiological Specifications for Foods, 2002. Microbiological Testing in Food Safety Management (Microorganisms in food 7). In: Chapter 2: Evaluating Risks and Establishing a Food Safety Objective. Kluwer Academic/Plenum Publishers, New York, pp. 23–43.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. In: Continuous Univariate Distributions, vol. 1. John Wiley and Sons, Inc.

Kanny, G., Moneret-Vautrin, DA., Flabbee, J., Beaudouin, E., Morisset, M., Thevenin, F., 2001. Population study of food allergy in France. J. Allergy Clin. Immunol. 108 (1), 133–140.

Kleinman, R.E., Bahna, S., Powell, G.F., Sampson, H.A., 1991. Use of infant formulas in infants with cow milk allergy – a review and recommendations. Pediatr. Allergy Immunol. 4, 146–155.

Kopper, R.A., Odum, N.J., Sen, M., Helm, R.M., Stanley, J.S., Burks, A.W., 2005. Peanut protein allergens: the effect of roasting on solubility and allergenicity. Int. Arch. Allergy Immunol. 136 (1), 16–22.

Kroes, R., Galli, C., Munro, I., Schilter, B., Würtzen, G., 2000. Threshold of toxicological concern for chemical substances present in the diet: a practical tool for assessing the need for toxicity testing. Food Chem. Toxicol. 38, 255–312.

Maleki, S.J., Hurlburt, B.K., 2004. Structural and functional alterations in major peanut allergens caused by thermal processing. J. AOAC Int. 87 (6), 1475–1479.

May, C.D., 1976. Objective clinical and laboratory studies of immediate hypersensitivity reactions to food in asthmatic children. J. Allergy Clin. Immunol. 58 (4), 500–515.

Mills, E.N., Valovirta, E., Madsen, C., Taylor, S.L., Vieths, S., Anklam, E., Baumgartner, S., Koch, P., Crevel, R.W., Frewer, L., 2004. Information provision for allergic consumers – where are we going with food allergen labelling? Allergy 59 (12), 1262–1268.

Ronteltap, A., van Schaik, J., Wensing, M., Rynja, F.J., Knulst, A.C., de Vries, J.H., 2004. Sensory testing of recipes masking peanut or hazelnut for double-blind placebo-controlled food challenges. Allergy 59, 457–460.

Sampson, H.A., 2004. Update on food allergy. J. Allergy Clinical Immunol. 113, 805–819.

Sicherer, S.H., Morrow, E.H., Sampson, H.A., 2000. Dose–response in double-blind, placebo-controlled oral food challenges in children with atopic dermatitis. J. Allergy Clin. Immunol. 105 (3), 582–586.

Sicherer, S.H., Noone, S.A., Munoz-Furlong, A., 2001. The impact of childhood food allergy on quality of life. Ann. Allergy Asthma Immunol. 201, 87 (6), 461–464.

Taylor, S.L., Hefle, S.L., 1997. The US experience. In: Lessof, M. (Ed.), Food Allergy Issues for the Food Industry. Leatherhead Food Research Association.

Taylor, S.L., Hefle, S.L., Bindslev-Jensen, C., Bock, S.A., Burks, A.W., Christie, L., Hill, D.J., Host, A., Hourihane, J.O'.B., Lack, G., Metcalfe, D.D., Moneret-Vautrin, D.A., Vadas, P.A., Rance, F., Skrypec, D.J., Trautman, T.A., Malmheden Yman, I., Zeiger, R.S., 2002. Factors affecting the determination of threshold doses for allergenic foods: How much is too much? J. Allergy Clin. Immunol. 109, 24–30.

Taylor, S.L., Hefle, S.L., Bindslev-Jensen, C., Atkins, F.M., Andre, C., Bruijnzeel-Koomen, C., Burks, A.W., Bush, R.K., Ebisawa, M., Eigenmann, P.A., Host, A., Hourihane, J.O., Isolauri, E., Hill, D.J., Knulst, A., Lack, G., Sampson, H.A., Moneret-Vautrin, D.A., Rance, F., Vadas, P.A., Yunginger, J.W., Zeiger, R.S., Salminen, J.W., Madsen, C., Abbott, P., 2004. A consensus protocol for the determination of the threshold doses for allergenic foods: how much is too much? Clin. Exp. Allergy 34 (5), 689–695.

Vieths, S., Schoning, B., Jankiewicz, A., 1993. Occurrence of IgE binding allergens during ripening of apple fruits. Food Agr. Immunol. 5 (2), 93–105.

Vieths, S., Jankiewicz, A., Schoning, B., Aulepp, H., 1994. Apple allergy: the IgE-binding potency of apple strains is related to the occurrence of the 18-kDa allergen. Allergy 49 (4), 262–271.

Vieths, S., Aulepp, H., Becker, W.-M., Buschmann, L., 1996. Characterization of labile and stable allergens in foods of plant origin. Food allergies and intolerances: proceedings of a symposium, Bonn, May 1995. VCH Publishers, Weinheim, pp. 130–149.

Vlieg-Boerstra, B.J., Bijleveld, C.M.A., van der Heide, S., Beusekamp, B.J., Wolt-Plompen, S.A.A., Kukler, J., Brinkman, J., Duiverman, E.J., Dubois, A.E.J., 2004. Development and validation of challenge materials for double-blind, placebo-controlled food challenges in children. J. Allergy Clin. Immunol. 113, 341–346.

Wensing, M., Penninks, A.H., Hefle, S., Koppelman, S.J., Bruijnzeel-Koomen, C.A.F.M., Knulst, A.C., 2002. The distribution of individual threshold doses elicting allergic reactions in a peanut allergic population. J. Allergy Clin. Immunol. 110, 915–920.

Zuberbier, T., Edenharter, G., Worm, M., Ehlers, I., Reimann, S., Hantke, T., Roehr, C.C., Bergmann, K.E., Niggemann, B., 2004. Prevalence of adverse reactions to food in Germany – a population study. Allergy 59, 338–345.