Probabilistic risk assessment model for allergens in food: sensitivity analysis of the minimum eliciting dose and food consumption

A.G. Kruizinga a, D. Briggs b,d, R.W.R. Crevel b,d, A.C. Knulst c,d, L.M.C. van den Bosch a, G.F. Houben a,d,*

a TNO Quality of Life, P.O. Box 360, 3700 AJ Zeist, The Netherlands
b Unilever Safety and Environmental Assurance Centre, Bedford, UK
c University Medical Center, Utrecht, The Netherlands
d Expert Group on Minimum Eliciting Doses of the European Branch of the International Life Sciences Institute (ILSI-Europe), Brussels, Belgium

Received 21 September 2006; accepted 26 September 2007

Abstract

Previously, TNO developed a probabilistic model to predict the likelihood of an allergic reaction, resulting in a quantitative assessment of the risk associated with unintended exposure to food allergens. The likelihood is estimated by including in the model the proportion of the population who is allergic, the proportion consuming the food and the amount consumed, the likelihood of the food containing an adventitious allergen and its concentration, and the minimum eliciting dose (MED) distribution for the allergen. In the present work a sensitivity analysis was performed to identify which parts of the model most influence the output.

A shift in the distribution of the MED reflecting a more potent allergen, and an increase in the proportion of the population consuming a food, increased the number of estimated allergic reactions considerably. In contrast, the number of estimated allergic reactions hardly changed when the MEDs were based on a more severe response, or when the amount of food consumed was increased.

Development of this work will help to generate a more accurate picture of the potential public health impact of allergens. It highlights areas where research is best focused, specifically the determination of minimum eliciting doses and understanding of the food choices of allergic individuals.

© 2008 ILSI Europe and TNO Quality of Life. Published by Elsevier Ltd. All rights reserved.

Keywords: Probabilistic model; Allergens; Food; Sensitivity analyses; Modelling

1. Introduction

To assess the risk arising from unintended exposure to food allergens a deterministic approach may be applied, where point estimates are used to calculate the outcome (Spanjersberg et al., 2007). In worst case scenarios usually the observed values of high percentiles, for example, the 97.5th percentiles or maximum are used. In less worst case scenarios mean values may be used. This approach often leads to conclusions such as “an allergic reaction cannot be excluded”, which is a poor basis for decision making in risk management. Furthermore, in this approach neither variability (for example, heterogeneity over different members of a population) nor uncertainty (for example, lack of knowledge about the true value) can be taken into account. In probabilistic modelling, distributions are used to represent the input parameters, making it possible to account for both variability and uncertainty and to quantify the risk.

A risk assessment model for allergens in food based on probabilistic techniques resulting in a quantitative
assessment of the risk and detailed information on the predicted number of allergic reactions has been developed at TNO (Spanjersberg et al., 2007). This approach is now recognized as a way forward in allergen risk assessment (Health Council of the Netherlands, 2007; Europrevall & FSA, 2007). In summary, in this model several input parameters can be identified, such as the proportion of the population who is allergic to a specific allergen, the likelihood that an allergic person consumes a certain product and if so in what amount, the likelihood that the food accidentally contains the allergen and if so the concentration of the allergen. These input parameters determine the distribution of intake of the allergen. By combining the distribution of allergen intake with the distribution of minimum eliciting doses (MED) among the population, the number of allergic reactions can be estimated more realistically.

In most probabilistic assessments, the majority of the variance in the output distribution is attributable to variability and/or uncertainty in a selected number of inputs. The identification of significant contributors to output variance has useful implications for future research. It makes it possible to target resources to a small number of important inputs, rather than spread the resources across the entire set of inputs (Cullen and Frey, 1999). Sensitivity analysis of a model is used to identify these key contributors to variability and uncertainty in model outputs by describing the change in output as a function of the change in the input variable. By comparing the magnitude of changes it is possible to identify those inputs which have the most influence on the output.

Validation of the model by a comparison of the predictions with observational data is not possible at this moment due to the lack of observational (incidence) data. Nothing is known about how the incidence of allergic reactions in a specific population relates to the consumption of a food containing an adventitious allergen.

This model has been developed to assess the risk associated with a given allergen concentration in a food. This means that an established allergen concentration (distribution) will be the starting input parameter. With these sensitivity analyses we wanted to investigate the major sources of variance and uncertainty affecting the risk associated with such given concentration (distribution). Therefore, we performed sensitivity analyses, involving different scenarios in which the input parameters of the MED part and the food consumption part of the model were varied.

2. Methods

As the basis for the sensitivity analyses, a case study was used for the probabilistic risk assessment model. The probabilistic model and case study have been described in more detail by Spanjersberg et al. (2007). The model consists of two elements: the distribution of allergen intake and the distribution of MEDs within a population. Within the model the values of these two factors relative to each other, and the shape of their distributions determine the probability of an allergic reaction. The reference model that was used as a basis for the sensitivity analyses calculated the frequency of an allergic reaction associated with the presence of hazelnut proteins in chocolate spread due to cross-contact at the production facilities. The mean number of allergic reactions calculated was approximately 5 per 10,000 persons at breakfast in the reference model (Spanjersberg et al., 2007). It should be noted that some worst case assumptions were made in our reference model. However, the risks calculated from these sensitivity analyses are only meant to determine the relative influence of the input parameters on the output and should not be interpreted as actual risk estimates for the population in question, due to the various assumptions made.

In the case study, data on food consumption were derived from the third Dutch National Food Consumption Survey (Anonymous, 1998), in which a representative sample of 6250 individuals recorded their food consumption for two consecutive days. For the modeling, data from male subjects older than 18 years were used, because males were found to be higher consumers of the food under study. IgE-mediated allergic reactions to food are generally accepted to occur within minutes to about 2 h after exposure. Therefore, we calculated the predicted risk of an allergic reaction using one eating occasion, since the allergic complaints are linked to single eating occasions. Breakfast and lunch were selected since these were found to be eating occasions with the highest consumption of the potentially contaminated food product, in this case chocolate spread (Anonymous, 1998). Table 1a describes the input parameters in the case study, the reference model, that were used to estimate the number of persons that would suffer from an allergic reaction as a result of unintended allergen consumption.

2.1. Description of input parameters identical in all models

Three factors were identical in all scenarios examined: the proportion of the population who are food allergic (the prevalence), the likelihood that a food accidentally contains an allergen and the distribution of the concentration of the allergen in the product (Spanjersberg et al., 2007). The concentration of the allergen was kept identical since this model was developed to assess the risk associated with unintended exposure to food allergens taking the allergen concentration as the starting point.

Table 1a

<table>
<thead>
<tr>
<th>The reference model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input parameters in reference model</td>
</tr>
<tr>
<td>Proportion of allergic persons in population</td>
</tr>
<tr>
<td>Proportion of population consuming chocolate spread at breakfast</td>
</tr>
<tr>
<td>Amount of chocolate spread eaten at breakfast (g)</td>
</tr>
<tr>
<td>Proportion of population consuming chocolate spread at lunch</td>
</tr>
<tr>
<td>Amount of chocolate spread eaten at lunch (g)</td>
</tr>
<tr>
<td>Likelihood product is contaminated</td>
</tr>
<tr>
<td>Concentration if contaminated (mg/g)</td>
</tr>
<tr>
<td>MED (mg)</td>
</tr>
</tbody>
</table>

Input parameters estimating the number of persons in which an allergic reaction would occur as a result of unintended allergen consumption.
As the prevalence of hazelnut allergy is not known, an estimate of the overall prevalence of food allergy was used, based on the self-reported prevalence of food allergy found in several European countries (Hupkens, 1999).

In the model the amount of allergen in a given product is represented by two parameters: the probability of the product accidentally containing hazelnut protein and, where appropriate, the distribution of the concentration of allergen in the product. The concentration of hazelnut proteins used in this study was measured in three different brands of chocolate spread, in which hazelnut was not an ingredient (Koppelman et al., 1999). Each brand of nominally hazelnut free chocolate spread contained unintended (adventitious) hazelnut protein, the amount ranging from 11 ppm to 752 ppm. Based on these results, the probability that the chocolate spread contains unintended hazelnut protein was assumed to be 100%. The concentration of hazelnut protein in the product was defined by a lognormal distribution (mean: 0.3 mg/g, standard deviation: 0.4 mg/g).

3. Sensitivity analyses

The input parameters that were investigated both for the food consumption part and the MED elements are presented in Table 1b. For each scenario the input variable that is changed is presented together with the change in the distribution of the input variable.

3.1. Sensitivity analyses on the MED

3.1.1. Using different model choices for describing the distribution of MEDs

To determine the impact of using different statistical models for describing the distribution of the MED, the lognormal, loglogistic and Weibull distributions were used. These parametric distributions were chosen based on a selection made by Crevel et al. (2007). These distributions were fitted using SAS System (version 8.2). The reference model used the raw data distribution of the unfitted data of the MEDs determined from a double-blind placebo-controlled food challenge with hazelnut in 29 patients (Wensing et al., 2002a).

3.1.2. Shift in the distribution of the MED to account for differences in the allergenic potency of different protein sources

We wanted to observe the effect of a change in MED independently of any change in the shape of the MED. To determine the impact of a shift in the MED distribution associated with differences in allergenic potency, the distribution of the MED was therefore shifted, while the shape of the distribution was retained based on the information about the distribution of the MED for hazelnut protein (Wensing et al., 2002a). The different shifts of the MED distribution for hazelnut protein (reference model) were based on the medians of the MED distributions fitted for egg protein (Bock et al., 1978), cows’ milk protein (Baehler et al., 1995) and peanut protein (Wensing et al., 2002b). The choices were made to demonstrate the shift in the distribution of the MED to account for differences in the allergenic potency using different protein sources for which real data exist. Due to this shift, MEDs could in theory become lower than zero; these MEDs were considered to be zero.

3.1.3. Variation in the distribution of the MED associated with the nature of the allergic reaction

The lowest dose which elicits one type of reaction may differ from the lowest dose which elicits another type of reaction, for instance, oropharyngeal symptoms (OS).
compared to gastrointestinal (GI) reactions. This applies on an individual patient basis as well as on a population basis. Therefore, MED distributions differ depending on the type of reaction used to define the MED. To mimic differences in the nature of the allergic reactions we compared the risk assessment outcome using a MED distribution based on OS and a MED distribution based on other (more severe) reactions than OS. The data to perform this sensitivity analysis were available from a double-blind placebo-controlled food challenge with peanut protein in 26 persons with peanut protein allergy (Wensing et al., 2002b). In this study, all persons (n = 26) showed an OS response first, while 16 persons also showed a more severe (more than OS) response following a higher dose. This was usually a reaction in the gastrointestinal tract.

3.2. Sensitivity analyses on food consumption

In the model the consumption of a food containing adventitious allergenic protein is represented by two parameters: the proportion of the population consuming the food under evaluation (in the reference case chocolate spread) and the amount of the food consumed per eating occasion. Needless to say that the proportion of the population and the amount eaten by these consumers will have an impact on the outcome of the risk assessment. We therefore questioned what the predicted risk would be had the protein been present in a more commonly eaten food that is also consumed in higher quantities. We selected white bread as such a food to assess what would happen to the calculated number of reactions if a particular cross-contact allergen occurred with the same pattern of distribution in a food which was consumed more widely and in larger amounts.

We therefore altered both the proportion of the population consuming a food and the distribution of the amount of food consumed based on the consumption of white bread. To explore the impact of each variable, the proportion of the population consuming the food and the amount of food consumed were also varied individually. These scenarios are presented in Table 3.

4. Probabilistic modelling

The modelling and the sensitivity analyses were performed using the software package @-RISK (Palisade Corporation, 1997). For this purpose, the parameters that were kept identical throughout all scenarios, as well as the experimentally-derived distributions described in the earlier sections that were subject to the sensitivity analyses, were used as input parameter.

Each model of one scenario was run for 20,000 iterations and 25 simulations, indicating that the model calculated the probability of an allergic reaction 20,000 times and this was repeated 25 times. In one iteration it is determined whether an allergic reaction occurred or not by sampling from the input distributions. By performing this 20,000 times a mean and standard deviation for the probability of an allergic reaction can be calculated. By repeating this process 25 times, an estimate of the variance around the mean and standard deviation of the probability of an allergic reaction is established.

5. Results

All scenarios were performed for two eating occasions (breakfast and lunch). However, the predicted risk at lunch was similar or lower than the predicted risk at breakfast and led to the same conclusions regarding the sensitivity analyses. Therefore, only the predicted risks at breakfast are shown in this paper.

5.1. Influence of distribution used to model MED data

In Table 2 the calculated number of allergic reactions for different model choices is presented. With the exception of the lognormal distribution, the input distribution generally had little influence on the predicted number of allergic reactions. The unfitted raw data, loglogistic or Weibull distributions all predict approximately the same number of reactions, whereas the lognormal distribution generally resulted in a much lower estimate of the number of allergic reactions.

<table>
<thead>
<tr>
<th>Allergen</th>
<th>Raw data</th>
<th>Lognormal</th>
<th>Loglogistic</th>
<th>Weibull</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peanut</td>
<td></td>
<td>7.1</td>
<td>5.7</td>
<td>8.6</td>
</tr>
<tr>
<td>Milk</td>
<td></td>
<td>156</td>
<td>146</td>
<td>156</td>
</tr>
<tr>
<td>Egg</td>
<td></td>
<td>250</td>
<td>219</td>
<td>258</td>
</tr>
<tr>
<td>Hazelnut</td>
<td></td>
<td>6.5</td>
<td>5.7</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Table 2 Results of sensitivity analyses on differences in model choice of distribution of the MED and a shift in MED to account for differences in allergenic potency, mean and standard deviation (SD), number of reactions per 10,000 persons in the total population

<table>
<thead>
<tr>
<th>MED used to account for differences in allergenic potency</th>
<th>Data used to determine shape of distribution</th>
<th>Median (mg)</th>
<th>Factor of shift (times compared to reference)</th>
<th>Number of allergic reactions at breakfast (per 10,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* NA = not applicable.
5.2. Influence of shifting the MED distribution to account for differences in allergenic potency

The results of a shift in the MED distribution to represent allergens of different potency are presented in Table 2. This shows that for two examples where the median of the MED distribution of hazelnut protein is shifted to a higher median dose (similar to milk or egg protein) the calculated number of allergic reactions is zero. However, when the median of the MED distribution of hazelnut protein is shifted to a lower median dose (similar to peanut protein) the calculated number of allergic reactions increases very significantly (from about 3 to 200 allergic reactions per 10,000 persons at breakfast).

5.3. Influence of differences in MED distribution associated with differences in the nature of the allergic reaction

The effect of basing the MED distribution on OS reaction or more severe reactions is presented in Table 3. When the MED distribution was based on OS the calculated number of allergic reactions after unintended consumption of hazelnut protein was 6.6 cases per 10,000 at breakfast. When the MED distribution was defined on more severe reactions the estimated number of allergic reactions was found to be 4.7 per 10,000 at breakfast.

5.4. Influence of changing the proportion of the population using a food and the amount of food consumed

The results in Table 4 show that if both the proportion of the population using a food and the amount of food consumed are adjusted to be comparable to those for white bread, the number of calculated allergic reactions was found to be approximately 22 times higher than in the reference model. To study the influence of each of the parameters separately, we adjusted either the proportion of the population using a food or the amount eaten. The proportion of the population using bread is approximately 17 times higher than the proportion using chocolate spread (35% instead of 2%). Increasing the proportion of the population using chocolate spread (35% instead of 2%) resulted in a proportionate increase in allergic reactions (68 instead of 4.6 per 10,000). However, increasing the amount consumed by approximately three times per eating occasion (57 grams instead of 19 g) resulted in a less than proportionate increase in the calculated number of adverse reactions (7.3 instead of 4.6 per 10,000). So, based on this simulation, the proportion of the population consuming the food has a greater influence on the number of predicted reactions than the amount consumed.

Table 3
Results of sensitivity analyses on the influence of differences in the MED distribution associated with the nature of the allergic reaction, mean and standard deviation (SD), number of reactions per 10,000 persons in the total population

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Number of allergic reactions at breakfast</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>MED based on OS of all persons</td>
<td>6.64</td>
</tr>
<tr>
<td>MED based on a more severe reaction</td>
<td>4.72</td>
</tr>
</tbody>
</table>

* To determine the MED distribution, data on peanut protein was used.

Table 4
Results of sensitivity analyses on the probability of and amount of consumption of potentially contaminated food products, mean and standard deviation (SD), number of reactions per 10,000 persons in the total population

<table>
<thead>
<tr>
<th>Data used to define the proportion of population consuming the food</th>
<th>Data used to define the amount of food consumed</th>
<th>Number of allergic reactions at breakfast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Chocolate spread</td>
<td>Chocolate spread (reference model)</td>
<td>4.60</td>
</tr>
<tr>
<td>White bread</td>
<td>White bread</td>
<td>102.76</td>
</tr>
<tr>
<td>Chocolate spread</td>
<td>White bread</td>
<td>7.32</td>
</tr>
<tr>
<td>White bread</td>
<td>Chocolate spread</td>
<td>68.48</td>
</tr>
</tbody>
</table>

6. Discussion

To determine the risk arising from unintended exposure to food allergens using a probabilistic model is a step forward from the deterministic approach, since a probabilistic model gives an estimate of the proportion of the population at risk of having an allergic reaction. Particularly in scenarios where a zero risk will not be attainable, the most accurate estimation of the (residual) risk and remaining uncertainties is crucial, and probabilistic quantitative risk assessment is a tool providing information on these aspects. However, decisions regarding the input parameters can influence the number of allergic reactions calculated. In this paper, the influence of variations in the median MED value and food consumption on the number of estimated allergic reactions was determined through sensitivity analyses. This work is a first step in determining the most important fields for generating new data to estimate more accurately the proportion of the population at risk of an allergic reaction due to a given concentration (distribution) of an adventitious allergen in food.

The choice of distribution used to represent the MED affects the shape of the distribution at low doses (Crevel et al., 2007). However, the effects of this parameter on the risk assessment have not been examined so far. Therefore, we explored the impact of varying the choice of statistical distribution used for modelling the MED distribution on the calculated number of allergic reactions in a population. With the exception of the lognormal distribution, using a different input distribution to describe the MED had little influence the number of allergic reactions. The lognormal is constrained to be symmetrical on a log scale,
while the Weibull and loglogistic are not constrained in this way. Therefore these two distributions are better able to fit to asymmetrical data. The lognormal therefore has to compromise between the fit at the top of the data and the fit at the bottom – depending on the shape of the data, this may have a large impact on the mean of the estimated lognormal distribution.

However, at present, no evidence supports the use of one particular distribution in stead of any other. Nevertheless, we can conclude that the three other statistical models (unfitted raw data distribution, Weibull and loglogistic) which gave approximately the same result, would represent a conservative approach in comparison to the lognormal distribution as these predicted a higher number of reactions. From a risk management point of view the use of one or more of these other, more conservative distributions might therefore be considered as a preferred approach until the choice of statistical distribution can be validated experimentally.

As would be expected, when the dose distribution is shifted to represent a more or less potent allergen, the number of predicted reactions increases or reduces. An increase [rightward shift] in the median of the MED distribution by a factor 10 decreases the predicted number of allergic reactions to zero. On the other hand, an approximate halving [leftward shift] of the median of the MED distribution increases the estimated number of allergic reactions more than 50 times. The scale of the dependence of the number of predicted reactions on the MED highlights the importance to risk assessment of improving knowledge about this key biological parameter, and in particular the shape of its distribution. In this study we only examined a shift in the whole distribution, leaving its shape unchanged. However, this variable could clearly have a considerable influence on the calculated risk as well.

This study also considered the influence of differences in the MED distribution associated with differences in the nature of the reaction.

The participants in the food challenge study (Wensing et al., 2002b) first reacted with OS and the majority also showed a more severe reaction, mainly in the gastrointestinal tract, following a provocation at a higher dose. There is so far no uniform way of defining severity of symptoms, but in general it is agreed that gastrointestinal symptoms are more severe than OS (Sampson, 2003). Therefore, these differences in the nature of the allergic reaction can serve as a proxy for the severity of the reaction. Although it would be very valuable for risk assessment to know the MED distribution for different effects, generating such data in a controlled study is extremely difficult, if not impossible, since it is unethical to expose patients to higher risks than necessary. This is recognized by guidelines produced by various expert groups for the conduct of double blind placebo controlled food challenges (DBPCFC), which state that objective or repeatable and lasting subjective symptoms are a reason for termination of the DBPCFC (Bindslev-Jensen et al., 2004; Taylor et al., 2004).

In this study, a MED distribution based on reactions other than OS decreased the mean number of allergic reactions by 30% compared to the calculated risk using a MED distribution based on OS. Thus, the impact was low compared to the other scenarios investigated in this work. It should be noted however that not all persons in the provocation study from which the data were drawn showed another response than OS. Assuming that these persons would eventually show another reaction than OS at a higher dose, decreased the predicted number of allergic reactions even further (data not shown).

We also examined the predicted number of reactions if the hazelnut protein were to be found in a food which is consumed in greater amounts and/or by a greater proportion of the population. As expected, an increase in the proportion of the population consuming the food caused a proportionate increase in the predicted number of allergic reactions. In contrast, an increase in the amount of food consumed per occasion had less influence. This illustrates that the number of users of a specific food has an important bearing on the risk assessment of unintended allergen intake. Food consumption data based on a short time period, for example 2 days, is known to underestimate the number of users of individual foods, specifically foods not consumed frequently. Statistical approaches are developed to deal with this issue (Dodd et al., 2006). In our models we assumed that food consumption of allergic persons is not different from that in the general population. This may not be realistic, but as yet, no data are available on food consumption patterns in specific allergic populations. As the probability of consuming a potentially contaminated food had a relatively strongly influence on the results, we recommend that food consumption surveys focused on allergic populations will be conducted in the future.

Inevitably a number of assumptions were made in this work, and only a limited set of scenarios was considered. The input parameters such as the probability of being an allergic person and the distribution of the concentration of the allergen in a food remained identical throughout all the scenarios.

In this work the proportion of people allergic to hazelnut was assumed to be the same as the overall prevalence of food allergy found in several European countries. In order to further refine the predictions, this work requires that prevalence of specific allergies is determined. The European Integrated Project EuroPrevall (FP6 Contract 51400, www.europrevall.org) will generate significant data on prevalence and MEDs for common allergens in Europe.

Although changing these parameters will influence the calculated number of allergic reactions, we do not expect them to influence the conclusions that can be drawn from the sensitivity analyses performed on the MED and the food consumption part of the model.

Overall it can be concluded that the location of the distribution of the MED and the proportion of the population consuming a food has a large influence on the number of
allergic reactions calculated using the risk assessment model for allergens. While this might be expected intuitively, the model begins to put quantitative measures to this observation, which can improve the setting of priorities for risk management.

This work should help provide an understanding of the major factors which influence the likely frequency of allergic reactions to adventitious protein. It also provides a framework which will permit better integration of new data in to allergen risk assessment. Development of this work should therefore help to generate more accurate information to feed into risk management decisions on food allergens, leading to better prioritization of risk reduction measures.

Acknowledgements

The authors would like to thank the members of Expert Group on the Determination of Eliciting Doses of the Food Allergy Task Force of ILSI Europe for their input. Furthermore, ILSI Europe and the Ministry of Health, Welfare and Sports of The Netherlands are acknowledged for financially supporting this research.

This work was commissioned jointly by the Food Allergy Task Force of the European branch of the International Life Sciences Institute (ILSI Europe) and the Ministry of Health, Welfare and Sports of The Netherlands. Industry members of the ILSI Europe Food Allergy Task Force are: Ajinomoto Europe, Barilla G&R F.Ili, Bayer CropScience Bioscience, Danisco, Groupe Danone, Africa, Nestlé, PepsiCo International, Royal Numico, Tate & Lyle Specialty Sweetners, Unilever, Wild Flavors Berlin.

References


