



Risk assessment and food allergy: the probabilistic model applied to allergens

M.Q.I. Spanjersberg, A.G. Kruizinga, M.A.J. Rennen, G.F. Houben *

TNO Quality of Life, Department Food and Chemical Risk Assessment, Utrechtseweg 48, P.O. Box 360, 3704 HE Zeist, Netherlands

Received 16 December 2005; accepted 14 July 2006

Abstract

In order to assess the risk of unintended exposure to food allergens, traditional deterministic risk assessment is usually applied, leading to inconsequential conclusions as ‘an allergic reaction cannot be excluded’. TNO therefore developed a quantitative risk assessment model for allergens based on probabilistic techniques resulting in a more exhaustive risk assessment and more detailed information. By now, this approach is recognized as the future approach in allergen risk assessment. A case study (hazelnut proteins in chocolate spread) is presented as a proof of concept.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Food allergy; Risk assessment; Probabilistic; Modelling

1. Introduction

IgE mediated food allergy, in which IgE antibodies are produced following exposure to specific food proteins, occurs in up to 4% of the total population, based on prevalence surveys (Hupkens, 1999; Sampson, 2004) and forms a significant public health problem. Since allergic persons avoid consumption of allergen containing foods, allergic reactions mostly occur to hidden or undeclared allergens in products. In order to assess the risk of unintended exposure to food allergens, the traditional deterministic risk assessment is usually applied, leading to inconsequential conclusions as ‘an allergic reaction cannot be excluded’. In addition, this conservative approach generally overesti-

mates the actual risk since it is based on worst-case values. Dr. G. Houben was the first to propose a modelling approach for allergens based on probabilistic techniques (Houben et al., 2002), which is now generally recognized as a solution to this issue. By now, significant work is done by Bindslev-Jensen et al. (2002) and Taylor et al. (2002, 2004) among others, to further explore and optimize allergen risk assessment.

This paper describes the probabilistic approach as proposed by TNO to be applied to food allergens. First, some general information is given about risk assessment and modelling of the occurrence of allergic reactions. The probabilistic approach is then described. Finally, some results of a case study (hazelnut proteins in chocolate spread) used for this model will be presented as a proof of concept. The outcome of this case study is only indicative and illustrates the model application; the goal of this article is to show the possibilities of probabilistic modelling to allergens rather than calculating exact risks.

2. Risk assessment

Risk assessment is an empirically based process that estimates the risk of adverse health effects from exposure of an

Abbreviations: ADI, acceptable daily intake; 1D/2D, 1/2-dimensional; DBPCFC, double-blind placebo-controlled food challenge; IgE, immunoglobuline E; *k*, number of runs (with respect to the used formulas); LOAEL, lowest observed adverse effect level; LOD, limit of detection; LOED, lowest observed eliciting dose; *n*, number of iterations (with respect to the used formulas); NOAEL, no observed adverse effect level; P95, 95th percentile; Rt, risk on an allergic reaction; SD, standard deviation; TDI, tolerable daily intake; *U*, uncertainty; *V*, variability.

* Corresponding author.

E-mail address: geert.houben@tno.nl (G.F. Houben).

individual or population to a chemical, physical or biological agent. The health risk assessment process usually involves four steps: hazard identification, hazard characterisation (dose–response assessment), exposure assessment and risk characterisation (IPCS, 1999).

The *hazard identification* (step 1) includes the assessment of the evidence for potential adverse effects in humans based on the available data on the nature of these potential effects and mode of action of the substance. The data show whether the substance poses a health hazard, and, if so, in which situations. The target organ(s) and critical effect(s) (the symptoms first observed by an increasing dose) are identified.

In the *hazard characterization* (step 2) the relationship between the exposure and the possibility, nature and severity of an adverse effect following the exposure is characterised. For most types of effects it is assumed that there is a minimum dose or concentration below which adverse effects will not occur: the threshold. To determine the threshold, different doses are tested, usually in laboratory animals, but sometimes also in humans. The highest tested dose without adverse effects is called the no-observed-adverse-effect level (NOAEL). Based on the NOAEL established in an experimental study a human limit value can be calculated, taking into account uncertainties and differences in experimental design. Uncertainties and differences are accounted for by uncertainty factors (e.g. for inter-species differences, intra-species variability and exposure duration). For some types of substances it is assumed that every level of exposure could result in adverse effects, in which case no threshold is assumed to exist. This for instance is assumed to apply for genotoxic carcinogens.

In food allergy, the real (population) threshold is often not known. Therefore, for allergens the lowest observed eliciting dose (LOED) is used, which is in fact similar to the lowest observed adverse effect level (LOAEL) for chemicals.

It is generally considered that a dose–severity relationship exists for food allergens and that subjects who react to very low doses are those most at risk of a severe reaction (Hourihane and Knulst, 2005).

The third step is the *exposure assessment*, which determines the nature and extent of the exposure to a substance under varying conditions. This step requires information on (1) the concentration of the hazardous substance in the sources, in this case foodstuff and (2) the total intake of foods containing the hazardous substance, or, alternatively, (3) measuring the intake indirectly by means of biomarkers. Uncertainty with respect to the exposure due to limited information and restrictions in the applied models and methods should be taken into consideration.

The fourth step, the *risk characterisation*, integrates the information derived from the hazard identification, hazard characterization and the exposure assessment. The exposure level from the exposure assessment step is divided by the NOAEL estimated in the dose–response assessment step to express the risk in relative terms. If this value (a

point estimate) is higher than 1 or close to 1, the conclusion will be that there might be a health risk.

3. Deterministic approach

Traditionally, risk characterisation is based on a deterministic approach, meaning that the calculated risk is based on a point estimate, usually the worst-case value for each input variable (i.e. worst-case NOAELs, assessment factors and exposure levels). This worst-case approach is intended to ensure that even the most sensitive part of the population is protected under all conditions, and therefore generally overestimates the health risk. In the case of food allergens, the maximum intake of a food may be multiplied by the maximum concentration of the allergen in this food. This results in the maximum estimate of the consumption of the allergen. If this consumption is higher than the lowest LOED found or threshold assumed, a possible reaction to the allergen cannot be ruled out.

For food allergens, it will be difficult or even impossible to set a threshold based on worst-case estimates, as any (detectable) exposure level may lead to the conclusion that an effect in a sensitive part of the population may occur.

Another disadvantage of the deterministic approach is that the magnitude of the potential problem is not known: there is no indication about the proportion of the population that might be at risk. This worst-case modelling can also be repeated with mean consumption, mean concentration and lowest LOED value. This might give a more realistic image of the problem, but the proportion of the population at risk is still unknown.

TNO has therefore researched whether the risk assessment model can be refined using probabilistic techniques in order to accurately define the potential health problem.

4. Probabilistic approach

In a probabilistic risk assessment model both variability and uncertainty in input variables can be taken into consideration (Slob and Pieters, 1998). Variability represents the true heterogeneity over time, space and different members of a population. Examples of variability are inter-individual variability in consumption and in sensitivity to for instance an allergen. Uncertainty is a lack of knowledge about the true value of the quantity. An example of uncertainty is associated with the limit of detection of an analytical method and the exploration of the threshold value outside the range of measurements. In contrast to the variability, uncertainty can be decreased, for example, by increasing the number of data points or by using a more accurate method of analysis. Both variability and uncertainty in hazard and in exposure data can be taken into account by using probability distributions to represent the input variables instead of point estimates. The data are plotted in a cumulative distribution curve. E.g. the LOED data and intake data can be plotted as probability distributions. By dividing the LOED distribution by the

intake distribution, the output distribution will describe the probability that a part of the population will be exposed at such levels and under such circumstances that adverse effects may occur. Consequently, an approximation of the percentage of the population likely to experience adverse effects at various exposure levels can be made.

It should be stressed that the probabilistic approach is applicable only to populations, and not to deriving safe intake levels for individuals. It is also more resource-consuming to select an appropriate model and perform the data analysis than using point estimates. However, the probabilistic approach makes better use of existing information, which facilitates more informed and better risk management decisions.

In terms of safety assessment, the probabilistic approach is applicable to the determination and validation of recommended limits of exposure. In terms of risk assessment, it is applicable for assessing the risk associated with exceeding recommended limits and assessing situations for which no limits have been established.

5. Probabilistic modelling of the relationship between frequency of allergic reactions and exposure

Fig. 1 shows the schematized presentation of the probabilistic approach in food allergy risk assessment. In this approach several input variables can be included: the chance that an allergic person consumes a certain product, in which amount this product is consumed, whether this food contains the allergen and in which concentration. These input variables determine the consumption of the allergen. Also, the LOED is determined. By comparing the LOED to the consumption of the allergen, it is determined whether an allergic reaction may occur and what the probability will be.

The schematic presentation clearly shows the key characteristic of the new risk assessment approach: where in worst-case modelling a point estimate (either a maximum or a mean) is chosen as the value for the input variable, in probabilistic modelling a data distribution is used.

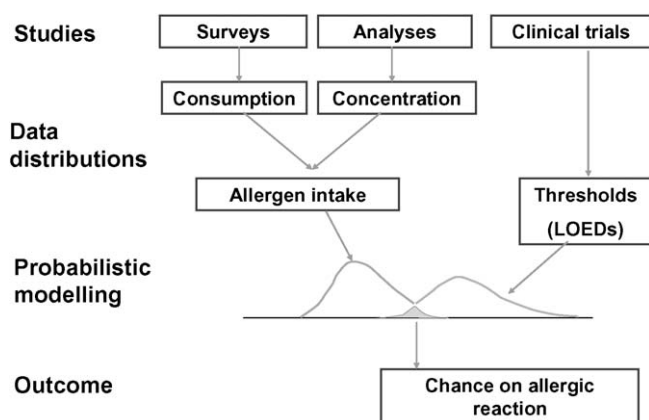


Fig. 1. Schematic presentation of the probabilistic approach in food allergen risk assessment.

6. Case study: an allergic reaction to hazelnut proteins in chocolate spread

Chocolate spread may contain hazelnut protein due to cross-contamination via production facilities and may thus form a threat to patients with hazelnut allergy. Allergic reactions due to the presence of these proteins in this type of product have been used as a case study for the probabilistic risk assessment model.

The four input variables required to run this model are the prevalence of hazelnut allergy, the range of LOEDs in hazelnut allergic patients, range of consumption of chocolate spreads and the range of hazelnut protein concentrations in this type of product.

First, the modelling part is explained.

6.1. Modelling

Distribution modelling is performed with specific software, like @-RISK, in which probabilistic models can be defined (Palisade Corporation, 1997). Within this software a value (an iteration) from the distribution used for the input variable is sampled repeatedly to calculate the result of the model, in this case whether or not an allergic reaction is to be expected. The number of iterations can, for example, be set at 1000. When calculating the model for 1000 iterations, 1000 data points are generated for which it is calculated whether or not an allergic reaction occurs. These data points give an estimate of the probability of an allergic reaction.

In the software the number of runs for each simulation can also be defined. The number of runs is the number of times the defined iterations are run, e.g. 10 runs of 1000 iterations. Each run gives an estimation of the probability of an allergic reaction. These runs together, the complete simulation, result in a mean and standard deviation of the probability of an allergic reaction (see Fig. 2).

For example, if in the first run in 2 of the 1000 iterations an allergic reaction is calculated to occur, the chance of an allergic reaction is $2/1000 = 0.002$. Ten runs results in 10 values for the chance of an allergic reaction. If the complete simulation (10 runs of 1000 iterations) is performed, the

$$R_t = \sum_{k=1}^k \frac{\# \text{allergic reactions in run } k}{n}$$

$$\begin{aligned} R_t &= \text{mean risk of allergic reaction} \\ k &= \# \text{ runs} \\ n &= \# \text{ iterations} \end{aligned}$$

$$SD(R) = \sqrt{\sum_{k=1}^k \frac{(R_k - R_t)^2}{k-1}}$$

Fig. 2. Mathematical formula.

Table 1
The five models used in the case study

Model number	Distributions of the input variables				1D/2D
	Prevalence of hazelnut allergy	LOED in hazelnut allergic patients ^a	Consumption of chocolate spreads ^a	Concentration of hazelnut protein in chocolate spreads ^a	
1	V	V	V	V	1D
2	$V + U$	V	V	V	2D
3	V	$V + U$	V	V	2D
4	V	V	$V + U$	V	2D
5	$V + U$	$V + U$	$V + U$	V	2D

V only variability is taken into account. $V + U$ both variability and uncertainty are taken into account.

^a Lognormal distributions.

chance of an allergic reaction is the mean of these 10 values.

Another important aspect in the software is the choice of the sampling method of the distribution. For real life situations Monte Carlo sampling, which randomly takes numbers from the distribution, is considered suitable (Gibney and Van der Voet, 2003; Monte Carlo project, 2003).

When variability alone is taken into account in defining the input distributions, the model is called one-dimensional (1D-model). Two dimensional models (2D-model) are models in which both variability and uncertainty are taken into account.

In this case study, we defined and used five different models (see Table 1).

6.2. Prevalence

At the time the case study was conducted, the prevalence of hazelnut allergy was not known with a sufficient degree of certainty. Therefore, in this case study a prevalence figure of food allergy in general was used. The 1D value we chose (5.4%) is a weighted average derived from several European prevalence studies (Hupkens, 1999). For the 2D-model a distribution, defined by minimum, most likely and maximum values, was chosen (1%, 3.7% and 5%, respectively).

As soon as data on hazelnut allergy prevalence becomes available, the outcome of this risk assessment can be refined.

6.3. Threshold (LOED)

In collaboration with the University Medical Centre Utrecht, TNO conducted a clinical study in order to determine the distribution of the LOED of hazelnut protein in a hazelnut-allergic population (Wensing et al., 2002). Twenty-nine patients with clinically established hazelnut protein allergy participated in a double-blind placebo-controlled food challenge (DBPCFC) performed with seven increasing doses of hazelnut protein (1 mg to 1 g). The challenge procedures were repeated for four persons. Four patients responded to 1 mg, nine to 3 mg, three to 10 mg, seven to 30 mg and six to 100 mg of hazelnut protein.

In this case study it is assumed that no difference exists in the allergic reaction between the raw and the roasted hazelnut allergen (the LOED was determined using raw hazelnut whereas in chocolate spread both raw and roasted hazelnut can be present). In reality, processing is known to unpredictably influence allergenicity of certain allergens. Yet, as this study was meant as a proof of concept and not to determine the actual risk, this aspect was not accounted for.

In the 1D-model, the variability between persons is taken into account, whereas the 2D-model also considered variability within persons.

6.4. Consumption pattern

In our model the consumption pattern is represented by two variables: the probability of using chocolate spread and the amount of chocolate spread used when consuming this product. The data are derived from the third Netherlands Food Consumption Survey (NL FCS from 1998 Hulshof et al., 2003), in which a representative sample of 6250 individuals record their food consumption for two consecutive days.

In this case study, we assumed that the pattern of food consumption of allergic persons does not differ from that of the general population. Furthermore, we based the chance of consumption of chocolate spread on one measurement day; i.e. both structural and incidental non-consumers of the product count in estimating the chance of zero-consumption. As this assumption may underestimate the actual risk of an allergic reaction, additional measurement days were evaluated in order to quantify this uncertainty in the chance of consumption.

In the 1D-model, the distribution of the amount of chocolate spread was used, whereas in the 2D-model the uncertainty in the chance of consumption is also taken into account. In this publication, we show the results of the models run for males only since this group consumes the most chocolate spread.

6.5. Concentrations

TNO determined the concentration of hazelnut proteins in three jars of different brands of nominally hazelnut-free

chocolate spread (Koppelman et al., 1999). The concentrations in the three brands were 752 ± 59 mg/kg, 115 ± 15 mg/kg and 11 ± 1.8 mg/kg (mean $\pm$ standard deviation). Allergen concentration is represented by two variables: the distribution of the concentrations of the allergen in the product (variability) and the probability of the product being contaminated with the allergen (uncertainty). For the 1D-model, the distribution of the concentration of hazelnut protein in chocolate spreads was chosen. The inclusion of an uncertainty in the concentration distribution was not performed for the following reason. A concentration above the limit of detection (LOD) represents a contamination, whereas a value below the LOD means that the product is either allergenfree or slightly contaminated. Since all three brands in our study were contaminated, no uncertainty exists about the presence of the allergen itself.

6.6. Software

In our case study we calculated the models using the software program @-RISK (version 3.5.2) and performed 25 runs of 20000 iterations for each simulation using Monte Carlo sampling.

6.7. Results

The results of our case study are presented below. First we show the results obtained by the traditional deterministic approach, using the mean allergen concentrations we found in the chocolate spread and the highest mean intake values (Table 2). The conclusion is 'an allergic reaction cannot be excluded' for brands 1 and 2. For brand 3, for which an intake of 0.2 mg was calculated, it is not possible to say whether an allergic reaction might occur, since the LOED found in our DBPCFC study was 1 mg, but lower doses were not tested. If worst-case values for allergen concentration and consumption are applied, the conclusion for brand 3 is 'an allergic reaction cannot be excluded' as well (Table 3).

If the probabilistic approach is applied, using data distributions, accurate information is obtained on the percent-

age patients at risk. Table 4 shows the mean and P95 chances of an allergic reaction for all five models at breakfast and lunch.

Consumption of chocolate spread, considering the 3 brands together, results in a highest mean risk of $<0.050\%$, meaning that less than 500 subjects per million persons will respond. At breakfast, consumption of chocolate spread products is higher compared to lunch, resulting in a higher risk of an allergic reaction at breakfast in the hazelnut-allergic population. For the 95th percentile, the highest risk at breakfast was found to be $<0.082\%$ (820 subjects in a million), compared to a highest P95 risk at lunch of $<0.049\%$ (490 subjects in a million).

The results of all models show that the risk of an allergic reaction at breakfast is higher than at lunch. This is obvious, as chocolate spread is mostly consumed at breakfast. The 1D-model (model 1) was also run for calculating the risk in women (data not shown), and resulted in lower mean risks compared to men. This is obvious as well, as men generally consume chocolate spread more often. The results of model 5 strongly resemble the values found in model 4, suggesting that the uncertainty in the consumption distribution influences the results the most.

For brand 3, which had been calculated separately (data not shown), a highest mean risk of $<0.004\%$ was found, meaning that less than 40 subjects per million persons will

Table 4

Results of the probabilistic approach (on the basis of the input variables described before)

Model number	Mean risk (in % (SD))		P95 (in %)	
	At breakfast	At lunch	At breakfast	At lunch
1	<0.043 (0.016)	<0.025 (0.010)	<0.065	<0.040
2	<0.037 (0.012)	<0.022 (0.010)	<0.055	<0.035
3	<0.050 (0.016)	<0.028 (0.011)	<0.082	<0.045
4	<0.045 (0.017)	<0.024 (0.010)	<0.069	<0.043
5	<0.042 (0.017)	<0.029 (0.012)	<0.068	<0.049

Note: the less than sign is used because some worst-case assumptions were made in the model (e.g. the prevalence of food allergy rather than the prevalence of hazelnut protein allergy was used).

Table 2

Results of the deterministic approach on the basis of mean concentration of allergen and highest mean intake of all consumption moments

Concentration of allergen (mean)	Consumption pattern (highest mean)	Intake (mg)	LOED	Risk on an allergic reaction
Brand 1: 0.752 ± 0.059 mg/g	19 g	14.3	<1 mg	Yes
Brand 2: 0.115 ± 0.015 mg/g		2.2		Yes
Brand 3: 0.011 ± 0.002 mg/g		0.2		Not known

Table 3

Results of the deterministic approach for brand 3 on the basis of worst-case concentration of allergen and maximal intake for all consumption moments

Concentration of allergen (mean + $3 \times$ SD)	Consumption pattern (maximum)	Intake	LOED	Risk on an allergic reaction
Brand 3: 0.017 mg/g	60 g	1.0 mg	<1 mg	Yes

respond. Thus, the lower contamination of brand 3 is clearly reflected in these results.

It should be noted that the results of this case study are only meant as a proof of concept and not as real risk figures, due to the various assumptions made.

7. Discussion

The TNO model is still experimental, but application of this probabilistic model gives an impression of the added value to allergen risk assessment. It shows the approach of a future way to assess the order of magnitude of the chance on or incidence of an allergic reaction associated with the presence of a certain range of allergen concentrations in a food product.

In the case study we described in this article, all types of allergic reactions were included, but the severity of the symptoms can also be included in the model by distinction of the threshold distribution into minor and severe reactions. In this way, the possible nature and/or severity of the reaction associated with the presence of a food allergen in a product can be predicted. This part of our investigation will be presented in a following publication.

A quantitative expression of risk permits appropriate levels of protection: the allergic part of the population can be protected against severe reactions, whereas unnecessary labelling or recalls may be avoided. The current precautionary 'may contain' labelling is based on the possible presence of an allergen rather than on the assessment of a quantitative risk. Since allergic persons avoid products on basis of the label declaration, precautionary labelling may unnecessarily limit the food choice of allergic consumers.

Not all required data are yet available (e.g. the threshold distributions of all known food allergens and the consumption pattern of the allergic population), but with increasing knowledge on allergic reactions and data distributions, the model can be refined time after time. Even additional factors that may influence the chance of an allergic reaction, for instance physical exercise, can be added to the model, under the condition that sufficient data are available.

In summary, TNO's quantitative allergen risk assessment model can be considered as an important strategy in improving allergen risk assessment and management approaches.

8. Conflict of interest statement

All authors state that there are none.

References

- 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.
- Gibney, M.J., Van der Voet, H., 2003. Introduction to the Monte Carlo project and the approach to the validation of probabilistic models of dietary exposure to selected food chemicals. *Food Addit. Contam.* 20 (suppl 1), S1–S7.
- Houben, G., Penninks, A., Brussaard, T. Changes in practice of risk assessment: from possibility to probability. In: TNO International Food Allergy Forum 15–16 April 2002, the Netherlands.
- Hourihane, J.O.B., Knulst, A.C., 2005. Thresholds of allergenic proteins in foods. *Toxicol. Appl. Pharmacol.* 207, S152–S156.
- Hulshof, K.F.A.M., Brussaard, J.H., Kruizinga, A.G., Telman, J., Lowik, M.R.H., 2003. Socioeconomic status, dietary intake and 10y trends: the Dutch National food consumption survey. *Eur. J. Clin. Nutr.* 57, 128–137.
- Hupkens, C. Prevalence of food allergies in Europe. TNO report, V99,1100. TNO, Zeist. 1999.
- International Program on Chemical Safety (IPCS) 1999. WHO. Geneva.
- Koppelman, S.J., Knulst, A.C., Koers, W.J., Penninks, A.H., Peppelman, H., Vlooswijk, R., Pigman, I., Van Duijn, G., Hessel, M., 1999. Comparison of different immunochemical methods for the detection and quantification of hazelnut proteins in food products. *J. Immun. Meth.* 229, 107–120.
- Monte Carlo project. Practical guidelines document 3. Guidelines for the input of data into probabilistic models of exposure to food chemicals. 2003. Available from: <<http://www.tchpc.tcd.ie/montecarlo/>>.
- Palisade Corporation. Guide to using @-RISK. 1997. Palisade Corporation.
- Sampson, H.A., 2004. Update on food allergy. *J. Allergy Clin. Immunol.* 113, 805–819.
- Slob, W., Pieters, M.N., 1998. A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: general framework. *Risk Anal.* 18 (6), 787–798.
- 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.
- Wensing, M. et al., 2002. The range of minimum provoking doses in hazelnut-allergic patients as determined by double blinded placebo controlled food challenges. *Clin. Exp. Allergy* 32 (12), 1757–1762.