How Can Effects of Processing be Factored into Hazard Identification and Risk Assessment of Novel Proteins? What are Potential Data Gaps?

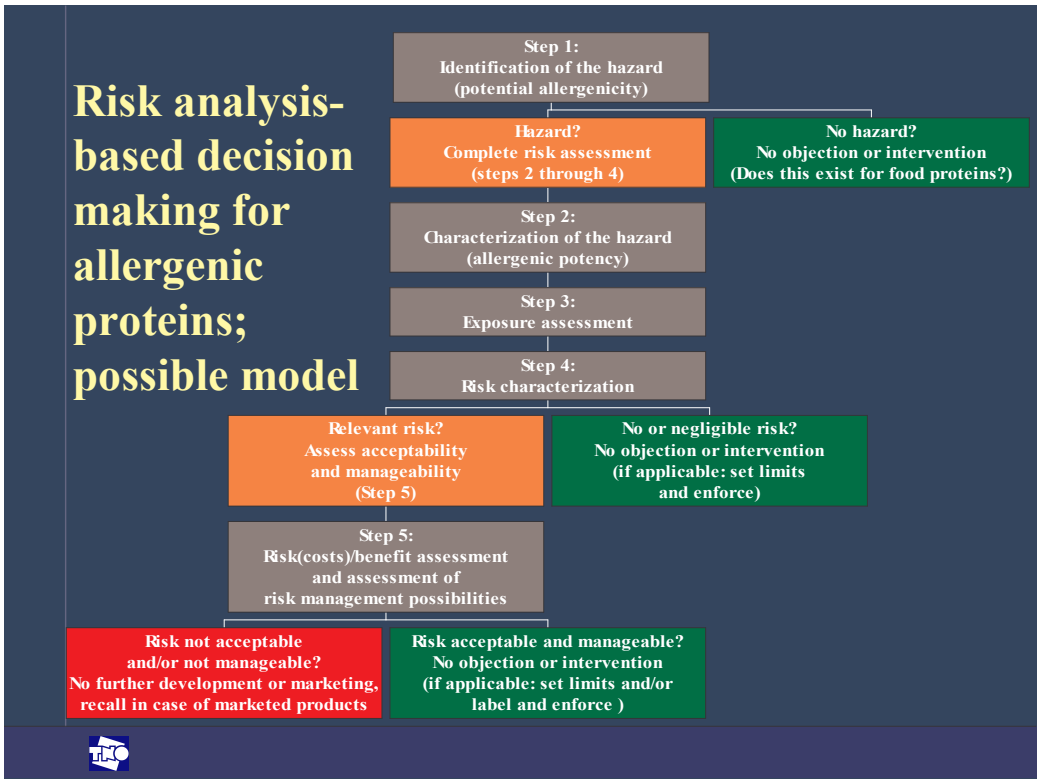
Geert F. Houben
TNO Quality of Life, Zeist,
The Netherlands

1. Introduction: risk assessment, risk management and risk analysis

During the past decades, more or less harmonised frameworks have been developed for risk assessment and risk management for substance in food. These have been incorporated into a risk analyses approach, in which the following steps are identified:

- hazard identification, in which the presence of a potential hazard is identified and the nature of the potential effects are assessed;
- hazard characterisation, in which the dose effect relationships for the effects are assessed;
- exposure assessment, in which the likely or potential exposure routes, levels and patterns are established;
- risk characterisation, in which the potential health risk is assessed on the basis of the information gathered in the previous steps;
- risk management, in which the anticipated health risk is weighed against other relevant aspects such as economical, technical or societal aspects, on the basis of which a risk management decision is made.

This basic approach nowadays is generally applied for chemical as well as microbiological hazards. In the mid nineties, TNO considered the applicability of this basic approach for dealing with risks of allergenicity of food proteins, and judged that there are no scientific reasons for not applying this same approach for food allergens. For allergenicity of food proteins, the risk analysis approach would include the elements outlined in the figure below.



In subsequent research projects during the last 10 years, TNO worked on methodologies and data generation for each of the steps outlined in the figure above, in order to develop this risk analyses based approach for allergenicity of food proteins. With respect to the questions to be addressed in this paper, the research and results in relation to steps 2, 3 and 4 are relevant. However, it is to be noted that the development of these steps by TNO was mostly focussed on the assessment of risks due to the presence of known allergens in food due to cross contamination in production facilities. Nevertheless, many of the methodological and data gap problems and solutions are more or less generic and are also relevant to the topic of this workshop. Therefore, in the next chapter, the risk assessment approach for food allergens is addressed in more detail. This is followed by a chapter addressing the questions of how effects of processing can be factored into hazard identification and risk assessment of novel proteins and what the potential data gaps are.

2. Risk assessment for allergens in food and use of available data

For developing a risk assessment approach for allergens in food, risk assessment approaches for chemical hazards in food were used as a starting point by TNO. Two different basic approaches are applied in risk assessment for chemical substances in food: one for effects for

which a threshold in toxicity is assumed to exist and one for effects for which a threshold in toxicity is not assumed to exist.

The approach for effects for which a threshold in toxicity is assumed to exist is applicable for most types of toxic effects. For these effects, the highest intake level of a chemical at which no adverse effect is known to occur in animals or humans (No-Observed-Adverse-Effect-Level or NOAEL) is established. From this NOAEL, a health based intake limit (e.g. an Acceptable-Daily-Intake or ADI) is derived by applying appropriate assessment and uncertainty or safety factors. An actual intake at or below this intake limit is considered to be safe. Such an intake limit can also be used to establish concentration limits for food products. For effects for which a threshold in toxicity is not assumed to exist, like for genotoxic carcinogenicity, this NOAEL/safety factor approach is not applicable. For genotoxic carcinogens, any exposure to the substance is assumed to result in an additional risk for cancer. For substances causing such effects, exposure should preferably be as low as reasonably achievable or practical. For practical reasons, a certain low additional cancer risk is often accepted (e.g. 1 additional case of cancer over a lifetime in a population of 10^6 persons), giving guidance for efforts in limiting the amounts of such substances in food and thus the exposure. The exposure level causing only such a low additional cancer risk is usually estimated by linear extrapolation on the basis of observed tumour incidences in animals or humans exposed at much higher doses.

Only limited information is available on thresholds with respect to allergenicity of food proteins. Yet, there are no scientific arguments to assume that such thresholds will not exist. This applies both for the sensitisation as well as for the effect elicitation phase of allergenicity. The existence of thresholds enables possibilities for a threshold approach in safety and risk assessment. However, additional research, method development, and data generation will be needed to establish the many individual or generic actual thresholds. Moreover, decisions will have to be made regarding the desired statistical confidence and level of protection.

Although a threshold approach may be applicable in certain situations, a threshold approach aimed at protection of all sensitive individuals would pose major practical problems. It is likely that such threshold approach would result in very low safe limits. This would in turn result in enormous efforts in developing practical analytical tools for monitoring and

enforcement of such low limits. Besides, compliance with limits aimed at the protection of all sensitive individuals would have major practical and economical consequences, that would pose major problems not only for industry, but also for consumers and regulators. It would in many cases practically lead to a zero-tolerance. In this respect, it should be considered that most, if not all, foreign proteins have an allergenic potency. As such, every protein in our daily diet may induce allergic reactions in some individuals. Yet, we accept the presence of allergens, including major allergens, in our daily diet and we even feed part of our offspring with one of the eight major sources of food allergens from the fist day after birth: cows milk. On the other hand, we know that about 90% of all documented food allergies world wide are due to only 8 main (groups of) allergen sources (i.e. peanut, soy, tree nuts, wheat, milk, egg, fish, and crustacea). This suggests that, apparently, many proteins in our daily diet pose no or only a minor concern. Allergenicity of (new) proteins or (new) protein containing products in our food supply therefore needs to be considered in a proper perspective.

A risk analysis based approach in which risks are weighed against other aspects might be a solution. This would however require the ability to assess the likelihood and nature and severity of the potential (rest) risks to be weighed against other aspects. This would ask for novel risk assessment approaches. A traditional risk assessment approach will (too) easily lead to a conclusion that "allergic reactions cannot be excluded" without providing further quantitative information. This can be illustrated on the basis of information on actual levels of protein allergens in food due to cross contamination in production facilities.

Traditional risk assessment for allergens in food

Several years ago, TNO performed a survey on the presence of hazelnut proteins in chocolate spread (Koppelman SJ, Knulst AC, Koers WJ, Penninks AH, Peppelman H, Vlooswijk R, Pigmans I, van Duijn G, Hessing M. Comparison of different immunochemical methods for the detection and quantification of hazelnut proteins in food products. Journal of Immunological Methods 1999; 229: 107-120). For 3 brands of chocolate spreads, for which the labels on the products did not list hazelnut (protein) as an ingredient, this resulted in the following concentrations (in mg hazelnut protein per g chocolate spread +/- SD). Brand 1: 0,752 +/- 0,059; brand 2: 0,115 +/- 0,015; brand 3: 0,011 +/- 0,002.

Subsequent investigation revealed that the presence of hazelnut protein in the products of these 3 brands was not according to recipe, but presumably due to cross contamination at the

production facilities. Since than, TNO has performed several surveys on residual proteins due to cross contamination. Also in these surveys, levels in the mg/kg range (generally between 1 and 5 mg/kg) were frequently observed. Most of the latter studies were conducted in facilities where cleaning regimes focussed on allergens were already in place, and the studies were meant to evaluate and improve the cleaning regimes. Based on this information, the contamination levels as found for the chocolate spread were considered realistic for situations were no specific cleaning procedures are in place. These levels therefore were used for a case simulation, addressing the question whether such levels would pose a risk for allergic consumers.

Besides research on hazelnut protein levels in chocolate spread, TNO in collaboration with the University Medical Centre of the Utrecht University in The Netherlands conducted a threshold study with hazelnut protein allergic subjects (Wensing M et al. The range of minimum provoking doses in hazelnut-allergic patients as determined by double blinded placebo controlled food challenges. Clin Exp Allergy 2002; 32(12): 1757-1762). In this study, 29 patients were provoked with increasing doses of hazelnut protein in a Double Blind Placebo Controlled Food Challenge (DBPCFC) protocol. Four of these 29 individuals already showed a response on the lowest tested dose of 1 mg protein. Finally, an analysis of the consumption of chocolate spread was performed on data from the 3rd Dutch National Food Consumption Survey (Hulshof KFAM, Brussaard JH, Kruizinga AG, Telman J, Lowik MRH. Socioeconomic status, dietary intake and 10y trends: the Dutch National Food Consumption Survey. Eur J Clin Nutr 2003; 57: 128-137). From this analysis it appeared that 8% of the Dutch population uses chocolate spread, with a highest mean consumption of 19 +/- 12 g (males during breakfast). The maximal consumption was found to be 60 g (males and females during breakfast and lunch). It should be noted that the hazelnut protein established in the chocolate spread will not have been of identical composition as the hazelnut protein used in the provocation study. A further variation is to be expected regarding possible hazelnut protein in the many consumer products that may be contaminated. However, assuming that the contamination levels as found for the chocolate spread are realistic for situations of cross contamination, above data can be used and combined in a risk assessment to assess whether contamination at such concentration ranges would be of health concern. Combination of the data in a traditional and worst case risk assessment leads to the following conclusions:

Contamination	Consumption	Intake	Risk?
Risk assessment based on mean point values			
Mean brand 1 (0.752 mg/g)	Highest mean (19 g)	> 1 mg (14.3 mg)	yes
Mean brand 2 (0.115 mg/g)	Highest mean g)	> 1 mg (2.2 mg)	yes
Mean brand 3 (0.011 mg/g)	Highest mean (19 g)	< 1 mg (0.2 mg)	?*
Risk assessment based on worst case point values			
Mean brand 3 + 3xSD** (0.017 mg/g)	Maximum (60 g)	> 1 mg (1.02 mg)	yes

* 4 subject responded at the lowest tested dose, and we do not know their threshold
** Mean + 3xSD is used as an estimation of the upper limit of the 95% confidence interval

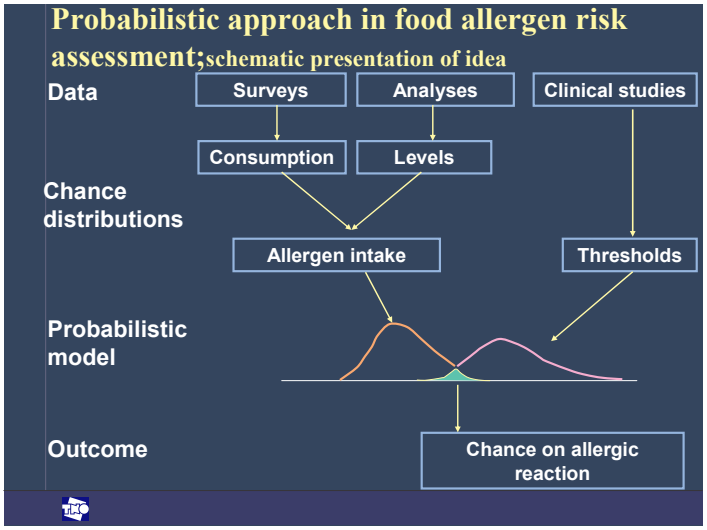
Based on the above risk assessment, an allergic reaction cannot be excluded for any of the concentration ranges considered. However, we do not have any information on the likelihood or frequency and nature and severity of possible reactions. To address this, TNO developed a risk assessment methodology for allergens in food, based on probabilistic techniques.

Probabilistic risk assessment for allergens in food

TNO aims at a probabilistic approach to finally describe the actual risk that we may encounter with various situations and scenarios. Besides for the assessment of situations where a limit set on the basis of a threshold approach is exceeded or where no such limits apply, probabilistic modelling may be used in the validation of established health based intake limits and to determine the statistical confidence and level of protection with such limits. The ultimate goal of probabilistic modelling in risk assessment for allergenic food proteins is to determine and describe the probability that an adverse event (of a certain type) should be expected to occur. This probability is determined by many chance determining factors, such as the chance of a person becoming (or having become) sensitised, the likelihood, level, and pattern of intake of products containing the allergen, the concentration pattern of the allergen in the food products, and various other circumstantial factors. In probabilistic modelling, it is tried to estimate the overall probability and statistical confidence

of a situation to occur by using information on probabilities or probability distributions for the various chance determining factors.

A simplified representation of probabilistic modelling is given in the figure below. In this figure, it is illustrated how information on thresholds for effect elicitation in sensitised individuals, information from food consumption surveys, and results from analyses on allergen levels in foods may be used to estimate the chance on an allergic reaction in a population of sensitised individuals.



Such a model can be refined and extended by using as much information as available. For instance, not all individuals that may consume the food product under evaluation will have an allergy, while only part of individuals with an allergy or allergies will suffer from food allergy. Again a smaller proportion of individuals will be sensitive for the specific food allergen that has contaminated the product under consideration. Examples of aspects that can be taken into account in the refinement and extension of models for probabilistic risk assessment for allergenic food proteins are given below.

- prevalence of food allergy (point estimate), either or not with distinction made between children and adults
- distribution of prevalences of allergen-specific food allergies, either or not with distinction made based on e.g. levels in foods
- physico-chemical and/or biological characteristics, like digestibility, stability in processing and preparation, effects in screening (e.g. animal) models

- distributions of thresholds for different allergens (from a worst case allergen assumption towards probability), either or not with distinction made based on screening results
- distributions of thresholds for different effects, e.g. based on severity or consequences
- consumption figures for subgroups of the population

TNO has developed various models for probabilistic assessment of the risks associated with allergens in food, and generally uses these models in parallel, to assess possible differences in predictions among the various models. The data assessed in the traditional risk assessment above were fed into these models, which resulted in the following risk estimates.

Results for the 3 concentration figures together:

- highest mean risk: < 0.05 % (< 500 x 10^{-6})
- P95 risk: < 0.082 % (breakfast) (< 820 x 10^{-6})
  < 0.049 % (lunch) (< 500 x 10^{-6})

Results for concentration figure of brand 3:

- highest mean risk : < 0.004% (< 40 x 10^{-6})
- P95 risk: < 0.02% (breakfast) (< 200 x 10^{-6})
  < 0.005% (lunch) (< 50 x 10^{-6})

Each figure above is based on the model that gave highest risk prediction (worst case choice), on all reaction types, and the total food allergy prevalence (in absence of reliable figures on the prevalence of hazelnut protein allergy).

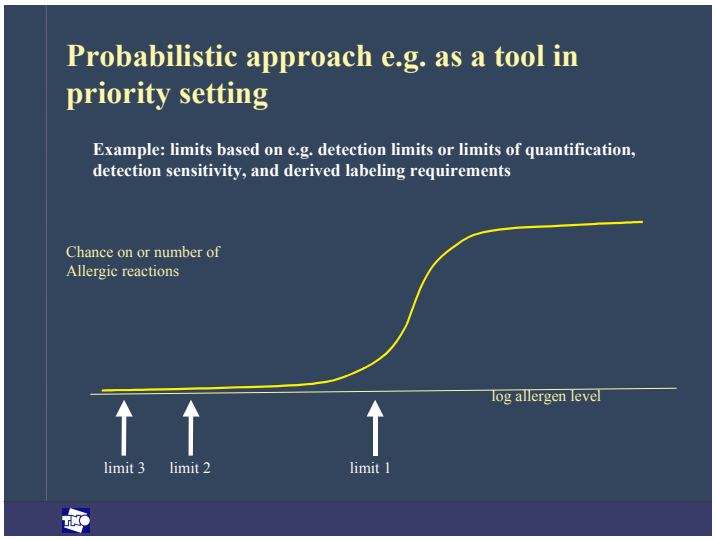
As the true incidence of allergic reactions due to cross contamination scenario's in a selected population is unknown, formal validation of the model is not possible with current data. However, it is possible to use sensitivity analysis to identify which parts of the model have the largest influence on the output (risk estimate). In addition, such a sensitivity analysis can also help to determine which variable(s) to focus on in future research to maximise the benefits of investments. Therefore, probabilistic sensitivity analyses were performed, involving different scenarios in which the input variables are changed. For each scenario only one input variable was varied. The assessment results from the initial analyses described above were used as a (relative) reference. Sensitivity analyses were performed both on the

Minimal Eliciting Dose (MED; threshold) part and on the exposure part of the model. We explored the impact of (1) using different statistical model choices for describing the distribution of MEDs, (2) a shift in the distribution of the MED to account for differences in the allergenic potency of different protein sources, (3) the likelihood that an allergic person consumes a certain product, (4) the amount of the food product consumed and (5) a variation in the distribution of the MED associated with the nature of the allergic reaction. 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 allergic reactions. In contrast, the number of allergic reactions hardly changed when the MED's 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 an allergen contamination. It highlights the areas where research is most valuable, specifically the determination of minimum eliciting doses and an understanding of the food choices of allergic individuals.

Probabilistic modelling in risk assessment for allergenic food proteins should be regarded as a step-by-step reduction in and/or description of uncertainties in the risk assessment. It is applicable already when information on the probability or probability distribution for one or several chance determining factors is available and it can be refined and extended continuously with new information becoming available. Although still with uncertainties, the ultimate goal of this approach is a better description of the risks of the presence of allergenic proteins in our food and thus a better starting point for risk comparisons, decision making, and risk management. The principle starting point is that not so much the possibility as such, but rather the probability of occurrence of an adverse event and its severity and consequences are to be considered and put into a proper perspective in the ultimate decision making and risk management.

It needs to be emphasised that the risk assessment method still is experimental. Nevertheless, application of this experimental model gives an impression of the added value that such an approach may have and that such approach may provide a (the?) future way to assess (the order of magnitude of) the change on or incidence of an allergic reaction and the possible nature and/or severity of such reaction associated with the presence of a certain range of allergen concentrations in a food product. This approach may also provide a way of assessing possible differences in (the order of magnitude of) risks (in terms of change on or incidence

of an allergic reaction) when a distinction is made based on the nature and/or severity of an allergic reaction. The probabilistic approach can also be used to assess possible differences in (the order of magnitude of) risks (now meant in terms of the change on or incidence of an allergic reaction and the possible nature and/or severity of such a reaction) when distinction is made between different allergens. Through sensitivity analyses insight can be gained into the major sources of variance or uncertainty, which may help us in assessing how progress may best be achieved. Could we best conduct (again?) another study on lowest eliciting doses or should we better investigate to what extend the consumption patterns of food allergic individuals resemble or differ from those of the general population? Or, to what extend may more sensitive analytical methods be of help…



In case an analytical method may enable monitoring or enforcement on limit 1, an increase in sensitivity towards limit 2 will result in a considerable contribution to the protection of consumers. However, a further increase towards limit 3 will not add significantly to consumer protection.

As such, probabilistic risk assessment and sensitivity analyses may serve as a tool in priority setting helping us to focus in spending our time and resources in our efforts for developing a reliable risk assessment. A critical scientific assessment and consensus regarding the basic risk assessment approach and models to be used and a further development of such approach and models are important future steps.

3. How can effects of processing be factored into hazard identification and risk assessment of novel proteins? What are potential data gaps?

In the previous chapter, examples of approaches for risk assessment for allergens have been described. It is also illustrated how available data can be factored into a risk assessment. In this, probabilistic approaches provide the possibility of making optimal use of all available

knowledge and data. The probabilistic approach developed by TNO was initially developed for the assessment of risks associated with the presence of allergens in food due to cross contamination in production facilities. However, the methodology in principle is also applicable for other purposes, as there is a high uniformity in the type of data needed and the steps involving a risk assessment. However, the approach as far as described in the previous chapter does not address all aspects needed for a hazard identification and risk assessment for all circumstances. For instance, the hazard identification step in the approach described was limited to the establishment of the presence of a known allergen in food, whereas for novel proteins or processed food products, establishment and characterisation of a possible sensitising potential of a protein also needs to be performed. This however was the topic of other lectures. For the ultimate risk assessment and data gap analyses, all these aspects need to come together. Below, a discussion table is drafted, as a first start in compiling an overview of available methodologies and data for risk assessment of proteins. This table may be extended, revised and refined on the basis of the outcomes of the workshop discussions and with developing scientific knowledge and information.

Overview of available methodologies and data for allergy risk assessment of proteins*

Step	Methodology available?	Can methodology be developed?	Data available?	Can data be generated?
Novel proteins				
Establishment of sensitising potential	Very limited	?	If possible, needed case by case	If possible, needed case by case
Dose effect assessment for sensitisation	Very limited	?	If possible, needed case by case	If possible, needed case by case
Establishment of population at risk for sensitisation	Very limited	?	No	?
Risk characterisation for sensitisation	No	?	No	?
For other steps: see known allergenic proteins				
Known allergenic proteins				
Identification of sensitised individuals	Yes	(improvement)	Yes	
Prevalence of sensitisation	Limited	Yes	Very limited	Yes

Step	Methodology available?	Can methodology be developed?	Data available?	Can data be generated?
Dose effect assessment for effect elicitation	Yes	(improvement)	Limited	Yes
Analytical identification of the allergen	Yes (approaches are available)	(improvement)	Limited (optimal methods not available for all allergens)	Yes
Analytical quantification of the allergen	Yes (approaches are available)	(improvement)	Limited (optimal methods not available for all allergens)	Yes
Food intake	Yes	(improvement)	Limited	Yes
Risk characterisation for effect elicitation	Limited	Yes	Limited	Yes
Processed proteins				
Establishment of sensitising potential	Existing epitopes: yes? New/altered epitopes: very limited	?	If possible, needed case by case	If possible, needed case by case
Dose effect assessment for sensitisation	Very limited	?	If possible, needed case by case	If possible, needed case by case
Establishment of population at risk for sensitisation	Very limited	?	No	?
Risk characterisation for sensitisation	No	?	No	?
Identification of sensitised individuals	Existing epitopes: yes New/altered epitopes: limited	Yes	If possible, needed case by case	If possible, needed case by case
Prevalence of sensitisation	Existing epitopes: limited New/altered epitopes: no	Yes	Very limited	Yes
Dose effect assessment for effect elicitation	Existing epitopes: yes New/altered epitopes: no	(improvement) ?	Existing epitopes: limited New/altered epitopes: no	Yes ?
Analytical identification of the allergen (once	Yes (approaches	(improvement)	Limited (optimal	Yes

Step	Methodology available?	Can methodology be developed?	Data available?	Can data be generated?
identified)	are available)		methods not available for all allergens)	
Analytical quantification of the allergen (once identified)	Yes (approaches are available)	(improvement)	Limited (optimal methods not available for all allergens)	Yes
Food intake	Yes	(improvement)	Limited	Yes
Risk characterisation for effect elicitation	No	?	Limited	?

* discussion table; this table may be revised on the basis of discussions and with developing knowledge and information