Appendix M: Validation Procedures for Quantitative Food Allergen ELISA Methods: Community Guidance and Best Practices

Although there are a number of documents published on method validation (1, 2) which target analytical methods in general, and there are numerous publications on validation of ELISA methods for pesticides, these documents do not address specific areas of concern for food allergen analysis, such as reference materials, spiking methods, or choice of matrixes. In the absence of a universally recognized reference standard for food allergen ELISAs, many organizations and end-users use different validation protocols and different analytical standards. Such inconsistency and duplication inevitably has a negative economic impact on the food allergen community. This document is designed to accompany the *AOAC Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis* (1), and to provide guidance specific to the validation of quantitative ELISA-based methods for food allergens. This protocol was designed to meet or exceed the minimum requirements set forth in the AOAC guidelines; it was developed with input from a wide range of experts in the area of food allergens, working under the auspices of the AOAC Presidential Task Force on Food Allergens and with the active contribution of the Allergen Working Group, part of the MoniQA network of excellence. This document will focus on developing guidance on a method validation study protocol to validate the performance characteristics of quantitative food allergen ELISA methods. The practical protocol is intended to help method developers in designing a study to generate appropriate validation data that would be suitable for submission to AOAC INTERNATIONAL or regulatory bodies for recognition. Both the study design and data would be subject to scrutiny before acceptance by the AOAC or other authority.

Methods for detecting various food allergens have been available for a number of years. Many of these methods use ELISA-based techniques to detect specific protein markers in food matrixes. The detection of food allergens by ELISA is a unique analytical procedure characterized by the recognition and binding of specific antigens by antibodies. Food allergens are proteins, which are large and complex molecules with defined structures in their native forms, that can induce allergic reactions in sensitized consumers. From the analytical point of view, the integrity of the protein structure is critical to favor protein solubility and promote antibody-allergen binding. Although specificity of antibodies in commercial ELISAs for food allergens varies, in most cases, these methods target a complex mixture of soluble allergenic and nonallergenic proteins, rather than a specific protein. This mixture of target proteins will have diverse structural and chemical properties in the complex mixture of a food matrix. Some food commodities contain several allergenic proteins, e.g., at least eight peanut proteins, such as Ara h 1 and Ara h 2, can potentially cause an immunological response. But other commodities, such as fish, shellfish, and mollusks, contain only one major allergen; still others may consist mainly of allergenic proteins, e.g., all major milk proteins (caseins, β-lactoglobulin, α-lactalbumin, etc.) possess an allergenic capacity.

The ability of an ELISA method to detect food allergen proteins in a test sample is affected by the efficiency with which these proteins are extracted from the sample, as well as the efficiency with which the antibody or antibodies used in the ELISA detect these proteins in the sample extract. The overall performance of an ELISA-based method for the detection of food allergens is a function of these two parameters.

The fact that allergic individuals often react to different protein constituents of the allergenic food further complicates the choice of targets. Because most food products are heat-treated, food production processes like roasting and extrusion can have significant influence on the solubility and extractability of the target proteins, as well as on the ability of the antibody or antibodies used in the ELISA to recognize them. Factors that may influence the test results include: (*1*) interactions with compounds in a food matrix (e.g., polyphenols and tannins); (*2*) reduced solubility and reactivity of heat-denatured proteins; and (*3*) differences in the protein profile of a particular food allergen from different species, varieties, and geographic origins. These factors all contribute to the difficulty in finding appropriate reference materials for food allergens and explaining why the proteins in a sample extract might not be fully comparable to that of the calibrators included with a particular detection method. These topics have been extensively reviewed recently (3).

Availability of validated methods is critical for both method developers and end-users. For method developers, validation of an analytical procedure is used to demonstrate that it is suitable for its

This document provides supplemental guidance on specifications for the development and implementation of studies to validate the performance characteristics of quantitative ELISA methods for the determination of food allergens. It is intended as a companion document to other existing publications on method validation. The guidance is divided into two sections: information to be provided by the method developer on various characteristics of the method, and implementation of a multilaboratory validation study. Certain criteria included in the guidance are allergen-specific. Two food allergens, egg and milk, are used to demonstrate the criteria guidance. These recommendations will be the basis of the harmonized validation protocol for any food allergen ELISA method, whether proprietary or nonproprietary, that will be submitted to AOAC and/or regulatory authorities or other bodies for status recognition. Regulatory authorities may have their own particular requirements for data packages in addition to the guidance in this document. Future work planned for the implementation and validation of this guidance will include guidance specific to other priority allergens.

These guidance and best practices were completed by the AOAC Food Allergens Analytical Community and submitted to AOAC INTERNATIONAL for publication in 2009.

Reference: Abbott, M., Hayward, S., Ross, W., Godefroy, S.B., Ulberth, F., Van Hengel, A.J., Roberts, J., Akiyama, H., Popping, B., Yeung, J.M., Wehling, P., Taylor, S., Poms, R.E., & Delahaut, P. (2010) *J. AOAC Int.* **93**, 442–450

© 2012 AOAC INTERNATIONAL

Table 1. Food commodities that should be included in cross-reactivity testing for ELISA methods targeting egg

Adzuki beans	Almond	Barley	Beef	Brazil nut
Buckwheat	Cashew	Chestnut	Chick peas	Chicken
Cocoa	Coconut	Corn	Crustacean/prawn/shrimp	Duck
Fish	Gelatin (bovine)	Hazelnut	Kidney beans	Kiwi
Lecithin	Lentils	Lima beans	Linseed	Macadamia nut
Milk	Oats	Octopus	Peanut	Peas
Pecans	Pine nut	Pistachio	Poppy seeds	Pork
Pumpkin seed	Rice—white and brown	Rye	Sesame	Soybean
Split peas	Sunflower seed	Turkey	Walnut	Wheat

intended purpose. For end-users, validated methods help to ensure reliability, repeatability, accuracy, and precision of the results generated using a particular method.

Method performance is documented using information and data provided by the method developer through interlaboratory validation studies. Minimum requirements for both information and data are included in this guidance, and may be applicable to any priority food allergen, as defined by the Codex Alimentarius Committee on Food Labeling in 1998 (4). However, due to the nature of food allergens, certain aspects, such as reference materials and spiking methods, would need to be addressed on a case-by-case basis. This document addresses these allergen-specific criteria for two food allergens, egg and milk. Further guidance for other priority allergens will be developed and communicated by the AOAC Presidential Task Force on Food Allergens and/or the Food Allergens Analytical Community under the auspices of the MoniQA network.

Required Allergen-Specific Information to be Provided on the ELISA Method

Information relating to the design of a method and its target analytes, as well as method performance characteristics, shall be provided by the method developer when submitting validation data for assessment. This information can be an important part of an overall package of information for evaluating a method. Proprietary information on antibody design or certain aspects of the method do not have to be disclosed. The AOAC guidelines (1) outline requirements for a final collaborative study manuscript. These allergen-specific requirements are additional recommendations that apply only to food allergen ELISA methods during method development and the final collaborative study.

The following information should be submitted along with the interlaboratory validation study data:

Antibody information.—Information on the antibody must include whether the antibody is monoclonal or polyclonal, whether it targets a single protein or multiple proteins, and whether the target protein used to generate the antibody was fractionated, modified, or synthesized in some way. Method developers are encouraged to include as much additional information about the antibody as possible. It is not necessary to reveal proprietary information. An example of antibody characterization for ELISA methods was discussed in a previous communication, specifically targeting mycotoxin/phycotoxin analysis (5). This approach could be adapted for allergen-specific antibodies.

Cross-reactivity.—Cross-reactivity is defined as a positive response to a sample that does not contain any of the target analyte. Method developers must test their allergen detection method for cross-reactivity for the target allergen in a variety of food commodities, which will vary for different target analytes and will depend on a number of factors. Food commodities tested for cross-reactivity should include a wide selection of foods and ingredients, particularly those that are genetically similar to the target allergenic commodity and that are likely to be analyzed for the presence of the target food allergen. The greater the number of items tested for cross-reactivity the better. In general, food items tested for cross-reactivity should be prepared as they would normally be consumed (raw or cooked).

Cross-reactivity testing should be based on the full-strength extracts, i.e., a sample of the item being tested for cross-reactivity should be extracted using the extraction buffer and procedure outlined in the method instructions, then analyzed at full strength to determine if it leads to a positive result. If a positive result is obtained, the extract must be diluted and rerun to characterize the extent of the cross-reactivity.

A minimum list of food commodities that should be included in cross-reactivity testing for egg and milk is provided in Tables 1 and 2, respectively. Many of these commodities will be the same for

Table 2. Food commodities that should be included in cross-reactivity testing for ELISA methods targeting milk

Almond	Barley	Brazil nut	Beef	Buckwheat
Cashew	Chick peas	Cocoa	Corn meal	Crustacean/prawn
Egg	Fish	Hazelnut	Lecithin	Lima bean
Oats	Peas	Peanut	Pecan	Pine nut
Pistachio	Poppy seed	Pumpkin seed	Rice–white and brown	Rye
Sesame seed	Soy bean	Split peas	Sunflower seed	Walnut
Wheat				

© 2012 AOAC INTERNATIONAL

Table 3. Matrixes of interest for ELISA methods targeting egg and milk

Egg	Milk
Chicken	Cookies, baked goods
Ice cream	Dark chocolate
Pasta	Drink mixes (ex. alcoholic beverage premix)
Salad dressing	Orange juice
Soy milk	Infant formula
Wine	Wine

all priority allergens, but specific items may be included on some lists, depending on particular concerns, e.g., genetic homology (crustaceans and dust mites) or matrixes of likely exposure. Table 3 lists matrixes of interest for ELISA methods that target egg and milk.

Information on calibrators.—The calibrators provided in the kit must be clearly defined. Information should address the following questions:

What is the calibrator that is supplied with the kit and used to generate the calibration curve? How was the calibrator prepared and assayed? Is the calibrator made from raw or processed material? Was the calibrator extracted or purified and if so how? Is the calibrator in extraction or dilution buffer?

It is very important to identify how the concentration of the calibrator is being expressed, what the units are, and whether it refers to the whole commodity or to a level of protein. If the calibrator is expressed as a level of protein, it should be clarified whether it refers to total protein or soluble protein and how the level of protein was determined, e.g., bicinchoninic acid assay with bovine serum albumin as the standard. Information on whether the calibrator is commercially available should also be provided.

Information on matrixes.—ELISA methods can be susceptible to matrix effects or perform differently in different matrixes. The method developer should clearly identify which matrixes the method is applicable for, on the basis of their in-house data, recognizing the variability of specific formulations. The developer should also identify any matrixes that the method is known to have difficulty with, and identify clearly which states of the food allergen (raw, cooked, or both) the method is capable of detecting.

LOQ, LOD, and lower limit of application (LLA).—LOD is defined as the lowest concentration or mass of analyte in a test sample that can be distinguished from a true blank sample at a specified probability level. LOQ is the lowest level of analyte in a test sample that can be reasonably quantified at a specified level of precision.

Manufacturers or method developers are free to define an LLA at whatever level of confidence they choose. This value may be higher than the LOQ and represents a level below which the method developer does not support or recommend use of the method.

Before conducting an interlaboratory study (precollaborative), a single-laboratory validation study of the ELISA-based allergen detection method should be carried out in-house by the method developer. Guidelines for single-laboratory validation of methods of analysis are readily available (2). The LOD should be estimated by a statistical analysis of the calibration data according to the ISO standard ISO 11843-2 (6) for linear data, or ISO 11843-5 (7) for linear and nonlinear data, using as default probabilities $\alpha = \beta = 0.05$, where α and β represent the probability of a false positive and false negative, respectively. When doing this estimation, care should be taken to include as many sources of variation as possible within a single laboratory. Calibration data from at least three analysts over a minimum of three different runs should be included, preferably using different instruments, if possible.

Ruggedness and lot-to-lot variability of method performance.—Ruggedness refers to the ability of a method to resist changes in the final results when minor deviations are made in the experimental conditions described in the procedure. The ruggedness of the method should be investigated by performing experiments in which specific parameters are changed to determine the impact on the experimental result. In particular, the effect of deviations in incubation times, reagent volumes, extraction conditions (time and temperature) should be investigated. It is recommended that deviations for time and volume be investigated at ±5% or more, and incubation temperatures tried at ±3°C or more. If any of these experimental conditions are particularly important in achieving consistent results, this should be mentioned in the kit insert information.

The shelf life should include the stability of all the reagents provided with the test kit, ideally through real-time testing of reagents under normal storage conditions. Accelerated stability testing at higher than normal storage temperatures can also be used to estimate stability. An expiration date for each test kit should be clearly indicated, along with appropriate conditions for storage before use.

A small number of test kits from each lot should be set aside for comparison with previous or future lots. When a new lot of test kits is produced, it should be tested against the previous lot. New lots should have characteristics similar to those of the previous lots. For example, a positive control sample, such as an incurred test sample or spiked sample, should be analyzed with each new lot to be sure that consistent results are achieved. Information on lot-to-lot variability should be provided by the kit manufacturer as part of the data submission package.

Key Elements of Interlaboratory Validation

Number of Laboratories Required

The required number of participating laboratories will be based on AOAC *Appendix D* guidelines (1), currently set at a minimum of eight laboratories contributing usable data at the end of the study.

In order to encourage participation from as diverse a group of laboratories as possible, the AOAC Presidential Task Force on Food Allergens and the Allergen Working Group of the MoniQA network require that, to minimize bias, no more than one-fourth of the total number of laboratories contributing data which is used in the final analysis of the study may be from the same organization. For the purposes of this requirement, the term organization refers to a particular company, such as the method developer or kit manufacturer, or to any other body, such as a regulatory body or other government agency.

Recruiting enough qualified laboratories to conduct a proper validation study for food allergens is difficult. However, the purpose of an interlaboratory validation study is to document the performance of the method in the hands of other laboratories, and this could not be accomplished if many of the laboratories participating in the study were from the same organization. If method developers use laboratories from their own organization as part of the validation study, the results generated by these laboratories shall have the same dispersion of results as those generated by other participating laboratories.

© 2012 AOAC INTERNATIONAL

Table 4. Example of raw data

	0 ppm		0.5 ppm		1.0 ppm		2.5 ppm		5 ppm	
Lab	A	B	A	B	A	B	A	B	A	B
1	0.61	0.46	1.10	1.13	1.24	1.97	3.08	2.80	3.65	3.61
2	–0.27	–0.41	0.41	0.29	0.57	0.71	2.80	2.07	4.51	4.84
3	0.37	0.21	0.62	0.11	0.45	0.70	2.82	2.93	4.24	3.93
4	0.13	0.13	1.06	0.62	0.79	0.41	1.95	2.37	5.22	4.96
5	0.24	–0.10	0.29	0.29	1.60	1.56	3.24	3.54	5.59	5.82
6	–0.23	–0.30	0.89	0.72	1.11	1.07	2.32	2.36	4.67	5.22
7	0.15	0.07	0.04	0.25	0.35	0.01	2.09	2.01	5.37	5.55
8	0.02	0.10	0.67	0.47	0.46	0.19	1.52	1.58	6.35	5.53
9	–0.02	–0.18	1.19	0.64	1.40	1.42	2.37	1.56	4.28	3.75
10	–0.10	–0.09	0.68	0.79	0.87	0.77	1.98	2.52	3.04	3.74

The AOAC Presidential Task Force on Food Allergens and the MoniQA food allergen community will attempt to develop a list of external laboratories from around the world that method developers could enlist to participate in validation studies. This will mitigate issues associated with the quality of results generated by the laboratories, or shipping of study samples across borders.

Number of Matrixes, Concentration Levels, and Replicates Required

The food allergen working group recommends that minimum requirements for any validation study include two matrixes, four concentration levels per matrix, and two replicate samples of each concentration per matrix in each laboratory. This is in compliance with AOAC *Appendix D* requirements for a minimum of five materials. For the concentration levels, one of the levels must be the zero level or blank. As an example, for a study using the minimum four concentration levels, two replicates and two matrixes, each participating laboratory would receive 16 samples for analysis.

In addition to a blank or zero level, one of the remaining concentration levels must be less than or equal to two times the LLA stated for the kit so that at least one of the concentration levels is at the lower end of the calibration curve. The remaining non-zero levels should be evenly distributed throughout the range of the calibration curve.

In general, more replicates per laboratory will result in greater certainty in the estimates of both repeatability and reproducibility. As with most estimates of variation, there is a law of diminishing returns with respect to increasing the sample size: the greatest advantage is made in the first few increases in sample size (replicates), but not much afterwards. These decisions are eventually made based on the tradeoffs between improved statistical estimates and resources needed to manage and perform the study. For allergen ELISA methods, the food allergen working group has concluded that a minimum of two replicates per laboratory will optimize the statistical confidence while not imposing undue burden on study participants.

Acceptance Criteria

Acceptance criteria are defined as numerical limits, ranges, or other suitable measures for acceptance of the analytical results to which a food allergen method should conform to be considered acceptable for its intended use. Acceptability of method performance is generally based on a number of factors, including percent recovery for spiked or incurred samples.

Ideal percent recovery levels would range from 80 to 120%. Recovery levels are affected by both the efficiency of the extraction step and the ELISA procedure. With ELISA methods for food allergens, this level of recovery is not always possible, particularly when certain difficult matrixes are analyzed. In addition, the recovery from incurred samples can be substantially different from those obtained using spiked samples. For this reason, recoveries between 50 and 150% will be considered acceptable so long as they can be shown to be consistent.

Data Analysis for Interlaboratory Studies

The ISO standard for method validation, ISO 5725-2 (8), and the AOAC *Official Methods of Analysis* (9) are the standards that outline how to analyze data stemming from interlaboratory trials in the context of analytical method validation. Each matrix/level combination should be treated as a separate experiment. For each matrix/level combination, the following analyses should be performed: Outliers should be tested sequentially by Cochran's and Grubbs' tests, as indicated in AOAC *Official Methods of Analysis, Appendix D* (1). Mean, accuracy (if applicable), repeatability (S_r), reproducibility (S_R), RSD of repeatability (RSD_r), and RSD of reproducibility (RSD_R) should be calculated and reported.

For each matrix, the LOD and LOQ of the method should be estimated using the sample S_R by the methods described in the IUPAC Nomenclature guidelines for LOD and LOQ (10). These guidelines call for a probabilistic estimation of LOD based on the variance observed at zero or near-zero concentration levels. If all assumptions are met (variance is constant and normally distributed, and the blank distribution is centered on zero), the LOD can be estimated as 3.3 times the SD of the distribution of blank results. This corresponds to false-positive and false-negative risks of 5% each ($\alpha = \beta = 0.05$), which is the recommended level for LOD estimation. LOQ can be set at 10 times the S_R.

Example of LOD Estimation for ELISA Collaborative Study Data

The following example uses data from a hypothetical collaborative study performed with an ELISA allergen test kit and shows the various steps required to calculate the LOD and LOQ for the method in a particular matrix as well as how to construct an operating characteristic (OC) curve for the method at a given concentration, such as the LOQ. Because different matrixes could give different results, data from each matrix in the study should be analyzed separately. The example is for samples spiked at nominal

© 2012 AOAC INTERNATIONAL

Table 5. Example of data analysis following AOAC/ISO 5725 Standard

		0 ppm	0.5 ppm	1.0 ppm	2.5 ppm	5 ppm
Total number of laboratories	p	10	10	10	10	10
Total number of replicates	Sum(n(L))	20	20	20	20	20
Overall mean of all data (grand mean)	$\bar{\bar{x}}$	0.040	0.612	0.882	2.395	4.694
Repeatability SD	s_r	0.108	0.211	0.220	0.305	0.325
Reproducibility SD	s_R	0.269	0.350	0.536	0.580	0.913
Repeatability RSD	RSD_r	273.438	34.456	24.888	12.721	6.925
Reproducibility RSD	RSD_R	680.549	57.203	60.711	24.228	19.455
HorRat value	HorRat	26.164	3.322	3.724	1.727	1.535

levels of 0, 0.5, 1.0, 2.5, and 5 ppm. The samples were analyzed in duplicate by 10 laboratories. It should be noted that these values may not reflect the full range of the calibration curve for this ELISA method, which could go much higher than 5 ppm. The results of the collaborative study and an example of how to use the data to calculate LOD are as follows:

Step 1: Collect data (*see* Table 4).

Step 2: Data analysis following AOAC/ISO 5725 standard (*see* Table 5).

Step 3: Model (S_R) by mean as per ISO 5725 (*see* Table 6).

Figure 1 gives an example plot of S_R versus mean. This model uses an ordinary least square estimate. Weighted least square analysis would also be acceptable.

Step 4: Estimate LOD and LOQ. Basic formula:

$$LOD = 3.3 \times s(0) = 1.0 \text{ ppm}$$

$$LOQ = 10 \times s(0) = 3.0 \text{ ppm}$$

Advanced formula to adjust for increase in s_R as mean increases: slope = 0.1285; intercept = 0.3081; xbar(0) = 0.039553; LOD = (xbar(0) + 3.3 × intercept)/(1–1.65 × slope); LOD = 1.3405; LOQ = 3 × LOD = 4.0215. These estimates are likely to be more accurate than those obtained following the simple formula.

Step 5: Construct OC curve based on results of Steps 3 and 4. Calculate the SD over a range of concentrations bracketing the LOQ using the formula:

$$SD = 0.1285 \times \text{concentration} + 0.3081$$

where 0.1285 and 0.3081 are the slope and intercept of the curve from Step 3.

Use a normal distribution calculation function to calculate the probability of obtaining a result higher than the LOQ (4.0) for the given concentration using the calculated SD and assuming a normal distribution. The probability thus calculated is plotted against the concentration to obtain the OC curve.

The curve below was calculated in Excel using the following equation to calculate the probability of a result higher than LOQ:

$$= 1 - \text{NORMDIST}(\text{LOQ, mean concentration}, S_R, 1)$$

where the LOQ is set at 4.0 ppm, the mean concentration is on the *x* axis, and the S_R is calculated from the mean concentration using the equation from Step 3.

Figure 2 presents an example of the OC curve. This OC curve shows the probability of obtaining a result above 4 ppm based on the concentration present in a sample. When the concentration in the sample is 4 ppm, there is a 50% chance the result will be above 4 ppm.

It is very important for collaborators to report all results obtained by the method without censoring to a predetermined LOD or LOQ. For nonspiked samples, this may mean half of the responses are negative numbers. It is critical to keep this information in the data set, as censoring will result in biased LOD/LOQ estimates.

For the results of the interlaboratory study, model S_R by concentration mean as detailed in ISO 5725-2. If the slope is significantly greater than zero, it should be taken that variance of the method increases with increased concentration. In this event, LOD estimates will need to be corrected with a general formula, which is shown above. If the general formula for LOD is used, LOQ can be estimated as three times LOD.

Additional guidance on the handling and analysis of data generated during interlaboratory studies will be provided through implementation studies conducted following this validation protocol.

Allergen-Specific Criteria

Certain criteria are dependent upon the specific target food allergen. For example, reference materials, spiking methods and food matrixes will vary from one food allergen to the next. General guidance on allergen-specific criteria and specific guidance for milk and egg allergens are as follows:

Reference materials.—Choosing a reference material for use in an allergen method validation can be extremely challenging. A perfect representative material rarely exists. Different species of the same food commodity may have different protein profiles. Processing methods can also drastically affect protein content, conformation, solubility, and reactivity. In general, a reference material is representative of the allergenic food commodity, is well-characterized, can be produced or supplied with robust reproducible

Table 6. Example of (S_R) modeling

Level	Mean	s_R
0	0.039553	0.26918
0.5	0.612395	0.350308
1.0	0.882414	0.535725
2.5	2.395355	0.580356
5.0	4.693936	0.913203

© 2012 AOAC INTERNATIONAL

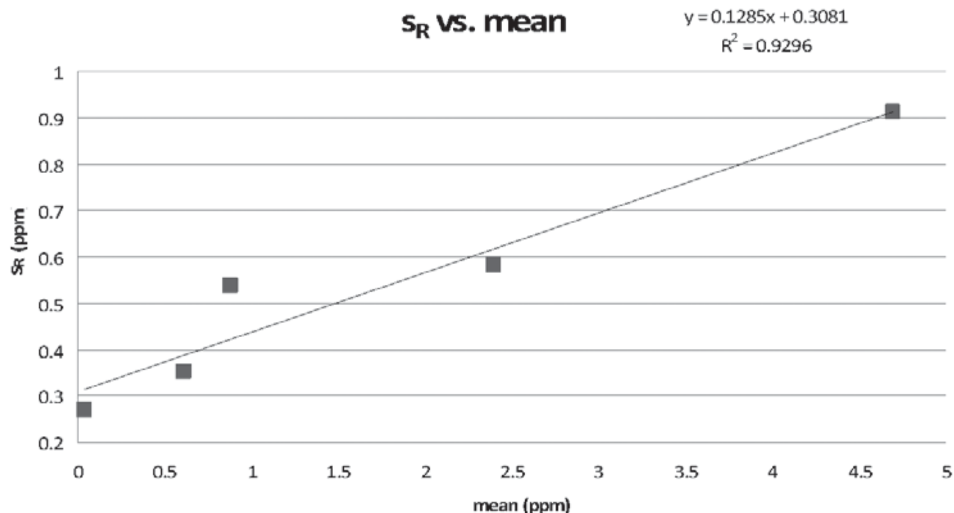

Figure 1. Example curve of S_R versus mean.

characteristics, and can be used as a calibration standard, control, or spiking material. Food allergens can be present in many different forms, processed or unprocessed, depending on the food matrix in which they are found, and with very divergent characteristics and functions in a food. It is unlikely a single material can represent many different possibilities at once. However, a widely available reference material will provide a common reference point for data comparison purposes between kits designed for the same food allergen.

For egg detection methods, based on a preliminary multilaboratory study, a suggested material is the National Institute of Standards and Technology (NIST) egg powder (NIST RM-8445). This is the first NIST reference material specifically intended for use in food allergen testing. The kit manufacturer is expected to provide a conversion factor relative to the NIST egg powder if a different material is used.

For milk detection methods, a suggested material is the NIST nonfat milk powder (NIST RM-1549). Although this reference material was not specifically intended for use in food allergen testing, it has been used in the past for method validations and has performed well as a reference material for milk ELISAs. The kit manufacturer is expected to provide a conversion factor relative to the NIST milk powder if a different material is used.

Spiking methods.—The best source of information on method performance for allergen detection methods is an incurred sample, which is defined as one in which a known amount of the food allergen has been incorporated during processing, mimicking as closely as possible the actual conditions under which the sample matrix would normally be manufactured. This kind of real-life sample would give the most accurate representation of the recovery and response of a particular method for that particular matrix. Whenever possible, validation studies for allergen detection tests should be run using incurred samples. Unfortunately, incurred samples can be difficult and costly to obtain, particularly in larger quantities required for a validation study.

Because of these limitations, validation studies using samples with food allergens added to them after manufacturing (spiked

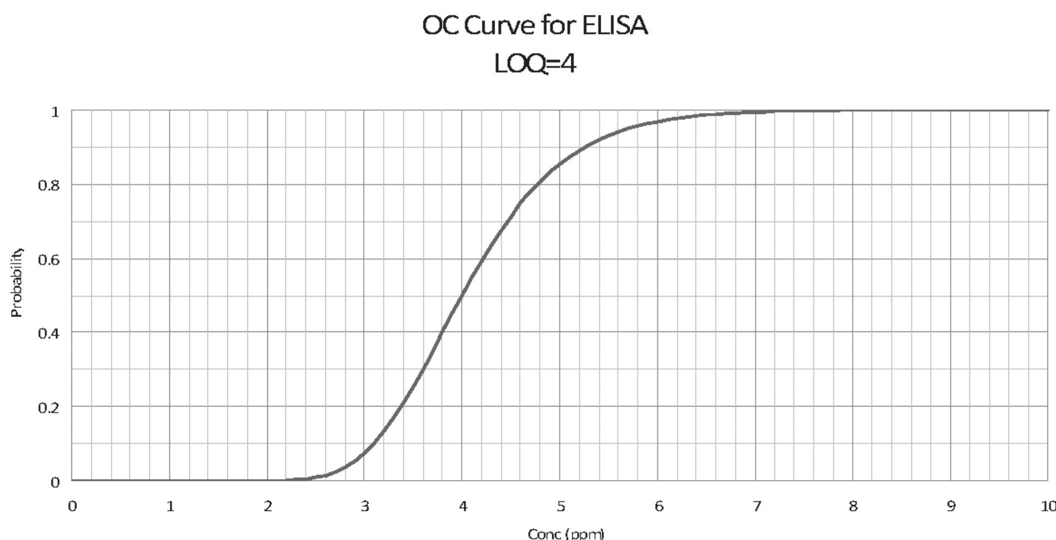

Figure 2. Example of OC curve.

© 2012 AOAC INTERNATIONAL

samples) are still considered an acceptable way to generate information about the performance of a method in specific matrixes. However, spiked samples may result in an artificially higher recovery than would be obtained from incurred samples; hence, some regulatory bodies may be unwilling to consider approval of validation data without the inclusion of data generated with incurred samples prepared with known and controlled amounts of the reference material for the allergen being targeted.

There are several ways to prepare spiked samples. One way involves the preparation of a large batch of a food sample that contains a specific food allergen, then gradual dilution of the allergen by mixing with more of the food sample that does not contain the allergen. This kind of sample preparation works best for samples that can be mixed well in an attempt to reach homogeneity, such as liquids or fine powders. An example would be the use of pasta containing a known amount of egg that had been ground to a fine powder and was then mixed with non-egg-containing pasta (also ground to a fine powder) stepwise until the desired concentration of egg was reached. Considerable effort is required to ensure sufficient mixing and to verify the homogeneity of the final batch of material, but this method of sample preparation has the advantage of being relatively similar to an incurred sample.

Because it can be difficult to mix a large batch of samples at a low spiking level to make a homogeneous mixture, the most precise way to spike samples is to add a known amount of a food allergen to each individual sample or test portion. This method results in each sample receiving an accurate amount of analyte, and addresses the issue of homogeneity of the spiked samples. Such a spiking method has been successfully used in the AOAC peanut *Performance Tested Method*SM study (11). In that study, individual test portions were weighed out and spiked before being sent out for analysis. This method of spiking results in a small part of the actual procedure (weighing of samples) being completed before the samples are distributed to study participants, and eliminates any weighing errors that may be introduced if study participants have to weigh the samples. Although this procedure is not ideal, the AOAC and MoniQA food allergens communities believe it is acceptable in order to overcome problems with production of large batches of food samples homogeneously spiked at a low level with a particular allergen. This type of sample preparation is the most artificial method and least representative of real-life samples.

When spiking samples, unaltered reference material should be used instead of a protein extract of the reference material. If the reference material is completely soluble in the buffer used for spiking, a solution of the reference material can be prepared and diluted to the appropriate level. The spike should be delivered in the same volume for each of the spiking levels.

The stability of the spiking material in the matrix of interest should be investigated by spiking several samples, and then extracting and analyzing them over the same period of time that will be required to complete the entire study. If the response changes significantly over time, this must be accounted for in the study design. Samples will have to be prepared, shipped, and analyzed within a defined time frame to avoid any decrease in response.

The suggested reference materials (NIST RM-1549 for milk and NIST RM-8445 for egg) are both powders that could be used with either of the spiking methods mentioned earlier (spiking a large batch of the matrix followed by serial dilution in a blank matrix, or spiking individual test portions using a spiking solution). Although the NIST nonfat milk powder (NIST RM-1549) is soluble in water or phosphate-buffered saline, the NIST egg powder (NIST RM-8445) is not. However, use of a tissue grinder, such as the Potter-Elvehjem type, will facilitate dispersion of the egg powder to form a homogeneous suspension. Thus, for both egg and milk, a stock solution of the reference material can be made, followed by dilution to the appropriate spiking levels. A recommended starting concentration for the stock solution is 1 mg/mL. In all cases, the method chosen for preparation of the spike and the spiking method should be documented in the validation report.

Food matrixes.—The matrix being analyzed can have a large impact on the performance of an ELISA method. Ideally, methods would be able to analyze all matrixes with equally reliable results. In reality, methods may work better for some matrixes than for others. The choice of matrixes included in a validation study is left to the method developer to meet customer demands. Although no matrixes are mandatory, some are of particular interest for each food allergen and are based on which food products are most likely to be contaminated with a particular allergen. Table 3 lists matrixes of interest for egg and milk. Method developers are encouraged to include as many of these matrixes as possible in their validation studies. However, good performance in one or even several matrixes does not guarantee good performance in others.

Conclusions

The food allergen analytical community is challenged to develop detection methods for multiple allergens in various food products to protect allergic consumers and promote consumer confidence. This protocol reflects the consensus reached through input from various food allergen analytical experts and contains recommendations based on the current knowledge of ELISA methods. Specific recommendations have only been included for two priority allergens, egg and milk. The general considerations of the protocol will be applied to other priority allergens in the future. Meeting the challenges of developing reliable food allergen detection methods requires conscientious and continuous support from the allergen community. Future work is planned for the implementation of this guidance document for egg and milk ELISA methods and for the development of similar guidance pertaining to other priority food allergens.

Acknowledgments

Under the auspices of the AOAC Presidential Task Force on Food Allergens and with the support of the MoniQA network of excellence (www.moniqa.org), the working group represents major food allergen test kit manufacturers and experts from regulatory agencies in Europe, Australia, Japan, Canada, and the United States, and the AOAC general referee for food allergens.

The participation of the following collaborators is greatly appreciated.

Mohamed Abouzied and Tony Lupo, Neogen Corp., Lansing, MI

Jeffrey R. Ammann and Armen Mirzoan, U.S. Alcohol Tobacco Tax and Trade Bureau, Beltsville, MD

Elizabeth Berryman, Marc Burke, and Richard Fielder, Tepnel Life Sciences, Flintshire, UK

Carmen Diaz-Amigo, U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, Laurel, MD

Valéry Dumont, CER Groupe–Laboratoire d'Hormonologie, Marloie, Belgium

Christiane Faeste, National Veterinary Institute, Oslo, Norway

Tatsuya Fujimura, Nippon Meat Packers, Inc., Tsukuba, Japan

Phil Goodwin, Diagnostic Innovations, Ltd, Denbighshire, UK

© 2012 AOAC INTERNATIONAL

Sigrid Haas-Lauterbach and Petra Urban, R-Biopharm AG, Darmstadt, Germany

Kurt Johnson, R-Biopharm Inc., Kalamazoo, MI

Terry Koerner, Food Directorate, Health Products and Food Branch, Health Canada, Ottawa, Canada

Steven M. Musser, U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, College Park, MD

Bruce Ritter, Elisa Technologies Inc., Gainesville, FL

Mike Ryan, Elisa Systems Pty, South Brisbane, Australia

Masahiro Shoji, Morinaga Institute of Biological Science, Inc., Yokohama, Japan

References

(1) *Official Methods of Analysis* (2002) AOAC INTERNATIONAL, Gaithersburg, MD, *Appendix D*

(2) Thompson, M., Ellison, S., & Wood, R. (2002) *Pure Appl. Chem*. **74**, 835–855

(3) Koppelman, S.J., & Hefle, S.L. (2006) *Detecting Allergens in Foods*, Woodland Publishing Ltd, Cambridge, UK

(4) FAO/WHO Codex Alimentarius Commission (1998) *Report of the 26th Session of the Codex Committee on Food Labeling*, ALINORM 99/22, May 26–29, Ottawa, Canada

(5) Fremy, J.M., & Usleber, E. (2003) *J. AOAC Int.* **86**, 868–871

(6) International Organization for Standardization (2000) *Capability of Detection, Part 2: Methodology in the Linear Calibration Case,* ISO 11843-2, Geneva, Switzerland

(7) International Organization for Standardization (2008) *Capability of Detection, Part 5: Methodology in the Linear and Non-Linear Calibration Cases*, ISO 11843-5, Geneva, Switzerland

(8) International Organization for Standardization (1994) *Accuracy (Trueness and Precision) of Measurement Methods and Results, Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method,* ISO 5725-2, Geneva, Switzerland

(9) *Official Methods of Analysis* (2005) 18th Ed., AOAC INTERNATIONAL, Gaithersburg, MD

(10) Currie, L. (1995) *Pure Appl. Chem.* **67**, 1699–1723

(11) Trucksess, M.W., Brewer, V.A., Williams, K.M., Westphal, C.D., & Heeres, J.T. (2004) *J. AOAC Int*. **87**, 424–428

© 2012 AOAC INTERNATIONAL