Review of allergen analytical testing methodologies: Overall Conclusions and Future Direction

10.1 Evidence Gaps

A number of gaps in knowledge and challenges to the allergen testing of foods have been highlighted in this report. They are summarised below:

As highlighted in the literature review, when tested by different kits, for some matrices cross-kit results are highly comparable. However, as reported by (Amnuaycheewa, Niemann et al. 2022 (in this case comparing gluten/gliadin- sensitive ELISA kits) while reporting similar results for some matrix types, yielded substantially different results for a few samples including samples of yeast extract, hemp protein powder and cookie. Those differences could be caused by any one of the several reasons: (a) differences in the grain source of glutens and related proteins, (b) differences in the efficiency of extraction and detection, (c) subsampling differences with particulates, or (d) some combination.

It is highlighted throughout this report that there is a lack of data and transparency regarding the applicability of test kits and public knowledge is often lacking regarding matrices for which kits have been found to be applicable and not applicable. This is demonstrated in the literature review for example by Röder, Kleiner et al. 2013 who state that there is a need to generate and publish data about the quantitative responses of methods to multiple cultivars of a given crop, and also to other allergenic foods.

Gaps in capability were also raised throughout the literature review relating to the inconsistency of data when a sample is processed compared to the testing of the raw product. This is perhaps best described by the detection of milk in processed products but applied to a wide range of allergens throughout the review. A crucial difference between kits designed to detect milk allergens is the detection of casein or casein and ?-LG compared to detection of total milk protein. In milk, the proteins with the greatest heat stability are casein and serum albumin and consequently kits which target these proteins rather than whey proteins demonstrate better detection of thermally processed milk (Nowak-Wegrzyn and Fiocchi 2009). In this work, the comparison of methods suggest that the target protein is crucial to method development with casein being less impacted by thermal processing in comparison to the total milk protein content. Additionally, the mass spectrometry method showed less variance between the cereal bar and muffin matrices and consequently may be more tolerant to a range of matrices, although there was not enough evidence in this paper to conclude this. As highlighted in our studies of PT data, for processed foods in particular, the less popular or available kits tend to be those most suited to denatured allergenic proteins. There is a fundamental lack of clarity in the public domain regarding the suitability of many test kits for more processed samples. This needs to be addressed to support accurate allergen testing.

The use of ELISA kits for the detection of milk in cheese was investigated by Ivens et al. (Ivens, Baumert et al. 2017(1), Ivens, Baumert et al. 2017(2)). The authors suggest that while milk residues are detectable in cheese by ELISA kits, it would be important to select the appropriate kit for the level of proteolysis the milk has undergone and they stress that it is possible that there are

fragments of proteolysed casein in processed milk which may not be detectable by modern ELISA kits.

Similarly, outcomes were reported in Section 7 which focussed on the study of proficiency testing data submitted to Fapas® between January 2018 and November 2022 relating to the variability of analysis between methods for the same allergen. Although regarding tree nut-allergens, much testing is progressing to LC-MS/MS methods, the majority of food testing laboratories world-wide continue to use ELISA kits which are specific not only to the allergen/matrix combination being tested but also specific to the design by the kit manufacturer. There was a fundamental lack of equivalence and reproducibility between different ELISA kit manufacturers (and sometimes between different types of kit for a single allergen by the same manufacturer). Batch-to-batch variation due to variations in kit-supplied buffers and calibration standards exists.

In addition to the issue of significantly different results being associated with different ELISA kit manufacturers, a further issue is evident in the PT data. This relates to the lack of comparability between ELISA kits from the same manufacturer. To provide one example, in PT 27316 (milk in infant soya formula) test material A, three kits from a single manufacturer were represented: at kit sensitive to ?-Lactoglobulin (R4912), a kit sensitive to Casein (R4612) and a kit for 'milk'. The consensus assigned values were, respectively, 1.56 mg/kg, 31.0 mg/kg and 17.8 mg/kg. Clearly, the assigned value for casein should not be nearly 2x that of milk. This issue is particularly prevalent in the PT data for egg and milk determinations across various kit manufacturers.

During Section 7, analysis of PT data clearly showed that the repeatability limit of ELISA kits has been reached. The mean ratio san/?p was only just being maintained, so the repeatability of ELISA kits had not improved during the course of the data analysis. The data also clearly showed that the reproducibility limit of ELISA kits had been reached. Separately, it was noted in Section 7 that the antibodies used to detect the allergen and the composition of calibration standards differs between manufacturers for kits detecting the same allergen, leading to variations in PT data depending on the method applied.

The potential to convert units of measurement in order to compare data between methods for the same allergen was also discussed in Section 7. The inter-conversion of units of measurement such as measurements of genomic sequence, peptide or single protein into a measurement of total food allergen is essential to evaluate and compare methods but conversion factors are often not supplied by kit manufacturers.

Throughout the literature review and the comparison of available test kits for each allergen (Appendix 1), there was limited data regarding how the LOD and LOQ were determined for each kit. Were they determined simply by analysing a buffer spiked with an allergen which would provide the lowest (i.e. most sensitive) LOD and LOQ data. Or were raw or processed samples used, spiked with allergen (which provided moderately challenging samples to interrogate). Or had the LOD and LOQ been determined by the most challenging and also most informative method, namely preparing a range of samples incurred with allergen and processed to a level representative of consumer foods. Due to the lack of transparency as to how LOD and LOQ was determined for each kit, the true comparison of kit capabilities is impossible for users taking the values published in the kit instructions. Instead, users would be required to prepare samples and compare kits experimentally, an exercise which is costly in time and finances. More details of this nature are required of manufacturers in order that users can make calculated decisions as to which kit or kits to pursue for their own intra-laboratory validations before they can offer testing services.

Given the data above regarding the suitability of different protein types depending on the level of processing of the samples, there is also a requirement that better transparency of the target protein of each kit would support allergen testing. There are issues here relating to Intellectual Property but the lack of this data for many test kits does not support food test. Instead, testing

laboratories must either rely on those kits for which the target protein is stated by the manufacturer, or perform costly studies themselves, comparing different kits and their suitability for different sample types.

All methods (ELISA, PCR and MS) suffer from issues in accurate quantification of allergens due to a lack of harmonised incurred reference methods, which is also a gap in testing capability.

A gap identified for allergen testing by LC-MS/MS is that a much more comprehensive knowledge of food genomes is required for mass spectrometry methods to work to their full potential in the analysis of food allergens, especially those from plant foods.

As highlighted by many stakeholders and our expert consultants, there is a need for the development and commercialisation of a fast, accurate, low-cost, multi-allergen tests, which can be used at point-of-use in factories to identify allergens, such as Near Infrared Spectroscopy or other innovative approaches. Relevant to the minimum sensitivity required for test kits, more work is also required to determine allergen thresholds of different foods for all allergens which must be declared on foods in the UK.

Finally, another evidence gap relating to allergen testing is this issue of the cross- reactivities which lead to false positive results when testing for some allergens. This is discussed below in more detail.

10.2 Cross-reactivity of testing methods

Cross-reactivities of individual test kits are detailed in Appendix 1. It is challenging to assess the full scope of cross-reactivities and whether there are kits on the market which compensate for the cross-reactivities in other kits for the same allergen due to the lack of cross-reactivity data provided by some kit manufacturers. A statement as to whether cross-reactivity testing has been conducted would be helpful, and a full list of those matrices tested would be ideal. However, cross-reactivities of note when considering the range of kits available for each allergen are:

As reported in the literature review, there are concerns regarding cross-reactivities including mustard antibodies cross-reacting with rapeseed (Brassica napus) at a level of 100%. This is a concern since rapeseed oil is used in a range of foods. .

Various milk testing kits report cross-reactivity goat and sheep milk. (Park, Coates et al. 2005) reported cross-reactivity between a peanut kit and chickpea, green pea, and lima bean.

There is much potential for cross-reactivity between crustacean allergens and insect food allergens, due to the commonality of certain proteins between the animal groups. In GB, interest in insects as a source of protein is growing. Insects that were legally marketed before 1 January 2018, submitted a novel foods application to the EU by 1 January 2019 and submit an application to the GB authorities by 31 December 2023, can remain in the market. Crustacean allergens such as tropomyosin are known to be highly homologous to their insect counterparts (Romero et al., 2016), and IgE cross-reactivity has been shown (De Marchi et al.,2021). It is therefore possible that antibody-based tests for crustaceans might detect insect proteins via species cross-reactivity.

Kamath et al. 2013 also reported increased recognition of multiple tropomyosin monoclonal antibodies upon heating of shellfish (Kamath, Rahman et al. 2013). These authors also reported cross-reactivity of tropomyosin in all 11 crustacean species, with partial detection in molluscs (cross-reacting with mussels, scallops and snails but not in oyster, octopus and squid). The authors conclude that specific monoclonal antibodies, targeting the N-terminal region of tropomyosin, must be developed to differentiate tropomyosins in crustaceans and molluscs.

Lacorn et al. illustrated in 2016 that closely-related plants show cross-reactivity to soybean ELISA (Lacorn, Dubois et al. 2016). However, although from a regulatory point of view, these cross-reactivities could be considered undesirable, they may still be relevant due to potential co-sensitivity amongst soy-sensitive consumers. Eighteen phylogenetically closely related species were tested. No cross-reactivities were observed for Lupinus angustifolius, L. albus, and L. luteus. In contrast, cross- reactivities were observed against Pisum sativum (dried and fresh seeds), Vicia pannonica, Lens culinaris, Arachis hypogea (roasted and raw), Cicer arietinum, Trigonella foenum-graecum, Trifolium pratense, Phaseolus vulgaris, P. lunatus, V. faba, P. coccineus, Vigna radiate, and V. angularis.

In the case of almonds, ELISA methods show cross-reactivities against phylogenetically closely related species including Prunus genus seeds (apricot, nectarine, cherry, plum, peach) but not with the flesh of these fruits (Slotwinski, Almy et al. 2018). Mahaleb cherry (Prunus mahaleb) is used as a spice. Apricot stones are used in the marzipan alternative persipan. Real-time PCR methods were developed. Burns et al. 2016 designed a real-time PCR method shown by the authors to be specific for Prunus mahaleb. Other work has also been completed using real-time PCR to distinguish almond and Prunus mahaleb to provide greater species specification compared to ELISA (Walker et al. 2018).

However, care must be taken to avoid cross-reactivity with Prunus mahaleb and using ELISA and PCR in combination can be an effective method of guarding against this.

As detailed further in Appendix 1, various of the nut kits show low levels of cross- reactivity with other nuts, and greater kit specificity is required. For example Sanchiz et al. (Sanchiz, Ballesteros et al. 2017) determined cross-reactivity of a pecan test kit with a food containing 5% of cashew nut, and chocolate with hazelnut, suggesting cross-reactivity with these two nuts.

Mass spectrometry can be more specific (less cross-reactivity) for detection of the target allergen protein to be quantified due to careful selection of the species-specific sequence to be detected, providing protein identity information, permitting a wider linear dynamic range, being less prone to be affected by food processing and, if appropriately applied, can be used as a reference method or for the production of CRMs. However, LC-MS/MS methods currently tend not to show the levels of sensitivity of ELISA methods and can also show low recovery, depending on the extraction method used.

The fact that celery can currently only be reliably determine by PCR is a concern, given the disadvantages of PCR data for allergen data interpretation.

As shown in Appendix 1, lupin ELISA kits cross-react with chickpea (0.0003%), lentil (0.0004%), soy flour (unroasted, 0.0700%), soy flour (roasted, 0.0009%) and soy lecithin (0.002%) although there are cross-sensitivities of these within allergen sufferers.

This cross-reactivity information, which is not always included in the manuals of testing kits, highlights the necessity for test kit manufacturers to perform a comprehensive range of testing to determine cross-reactivity, and to make this list public, to support allergen testing and management. Overall, a lack of transparency regarding the validation studies conducted (or not conducted) on each test kit is a fundamental limitation when it comes to determining the capabilities and limitations of test kits.

10.3 Overall Conclusions

This review of methods used for food allergen testing involved a literature review of data (laboratory evaluations and data from kit manufacturers) published in the public domain in this

field over the last twenty years (included in Sections 2 and 3). One issue which arose time and again was the concern that little or no data about which version of the various commercial testing kits are detailed in the literature. Commercial food testing kits often undergo development and alteration which will change the performance of the kits. For example, ELISA kits necessarily undergo alterations, when new batches of antibody are prepared when stocks are depleted. Also, kit manufacturers often update each of their kits, for example introducing altered or improved extraction buffers or other reagents, and performing more validation studies and, on occasion, the name and reference number of the kit is not altered. Much caution is therefore required when relating current kit performance to data published in previous years. The protein target of the kits is not always known (this is often the case when kits are underpinned by polyclonal antibodies which were raised against the allergenic food as a whole, so the precise protein/epitope is not known) or the protein identity is not disclosed due to proprietary issues. This lack of knowledge regarding the target protein poses an issue during incident management when application of range of tests which are sensitive to different protein targets would be beneficial to gain confidence when trying to detect an allergen in a sample or to rule out a suspect sample for allergen presence. The crucial challenge in ELISA detection is the variability of data generated across different kits targeting the same allergen. Another gap in the information included in many test kits is a lack of interconversion factors to convert from one allergenic protein to another or to ‘total allergen protein per kg food.’

The performance of allergen detection methods is affected, often but not exclusively, in a negative manner, when foods undergo thermal processing, a very common treatment of most food products unless commonly eaten raw. Thermal processing can, for example, alter protein folding and oxidation processes can take place which can lead to modification of amino acids, formation of protein bound carbonyls and aggregation. Each of these changes can affect the detectability of the parts of the protein which are detected in methods such as ELISA and can also alter the level of protein extractability/solubility which affects the recovery of testing procedures. Limitations involving detection of milk in cheese due to proteolysis during processing were also discussed, as were cross-reactivities which can lead to false positive data. Kit manufacturers should release their entire validation reports as a matter of course, and manuals should clearly state the food matrices for which their kit has been validated, whether these products were raw or processed and the form (e.g. time, temperature, pressure) of processing.

To help to overcome complications linked to determining any changes in recovery (and therefore LOD) linked to food processing, it would also be beneficial if it was stipulated that methods must be tested against standardised RMs (where available), ideally indicating which RMs should be used for each allergen and the disclosure of the RM used should be included in the validation data. This way, test users could compare the performance of the kit in their hands on the same RM compared to the performance declared by manufacturers. Ideally, harmonisation of methods would be achieved, and CRMs would be available for all allergens and for a much wider range of levels of processing, in order that users can reliably calibrate their methods against these RMs. Preferably, more resource would be allocated to preparing a much wider and comprehensive range of CRMs for each allergen. Currently, CRMs are lacking (incurred RMs in particular) with certain of those previously prepared now removed from the market. RMs are expensive to prepare and verify as homogenous and commercial incentive for their preparation is often low, hence their poor representation on the market. Realistically, without RMs against which to calibrate methods, the capability of current methods is only semi-quantitative. It should also be encouraged that kit manufacturers provide detailed information regarding the manner by which the amount of allergenic protein present in a sample can be calculated from the data per mass of food product, in line with the conclusions of the FAO/WHO expert consultation. As is currently the focus of FAO/WHO, more data are also required to determine allergen thresholds of different foods. These data will be valuable in terms of allergen management in the UK.

The review has determined that allergen test kit suppliers are often not transparent in publishing the entire validation studies completed before marketing the kit. A likely benefit of full publication

of validation results in the public domain would be improved protection of consumers. This would include how the test materials used for measuring kit performance were prepared. For example, were they simply a buffer spiked with the allergen - this comprises the simplest form of test material and is not representative of ‘real’ food samples. Was a raw food used and spiked or was a finished food product spiked with the allergen? Were incurred products prepared, accurately mimicking food processing techniques? Spiking approaches for test method validation are not representative of true food samples as the allergenic food ingredient used for spiking had not undergone the same level of processing as the food matrix. The use of incurred materials is the ideal scenario to mimic ‘real’ foods. As the choice of analytical method is crucial to ensure consumer safety, we need to have confidence in the testing methods used and standardisation to compare data. This relates to all allergens but soyabean is one key area where detection of the allergen is kit- or method-dependent, often relating to the level and type of processing. So many foods contain soyabean and each involves a different level of processing. This is discussed in this report (Section 3 and also Section 10).

To further complicate our understanding of the range of commercial testing kits available, it has become apparent during the preparation of Table 1 that not all suppliers detail much information regarding the performance of the kit. There is evidence from comparing some of the kit parameters that some suppliers re-brand other kits and sell under a different name, which further complicates our understanding of current capabilities and the range/number of different kits available on the market.

In order to protect consumers against accidental consumption of allergenic foods and associated allergic reactions, there is a need to improve education across the supply chain relating to the need for testing at each stage in the chain. It is also important that suppliers of foods and their ingredients understand the limitations of each particular test (Sections 3 and 4). The need for improved harmonisation of methods and auditing throughout the supply chain was also covered in these sections and legislative gaps highlighted. Alignment of food testing methods developed in the EFSA ThRAll project was also discussed.

This project also comprised statistical analysis of the Fapas® allergen proficiency testing data generated over the last five years. This analysis (Section 7) demonstrated that the repeatability and reproducibility limits of ELISA kits have been reached. New methods are clearly required to fill this gap. Furthermore, analysis of the Fapas® data reinforced issues raised by other authors that ELISA kits do not report like-for-like data and that data cannot be inter-converted. Finally, this statistical analysis exercise of Fapas® data raised the issue of significantly different results being associated with different ELISA kit manufacturers. A further issue is evident in this data, and it relates to the lack of comparability between ELISA kits from the same manufacturer, demonstrating that the use of different kits from one provider for a single allergen yields a different final level of allergen in a given sample.

It has been clearly demonstrated throughout this project that there are limitations in allergenic food ingredient detection and quantification. Each of the main testing methodologies show advantages and disadvantages and hybrid testing will yield higher efficiency in successful testing in most cases. Reliance on a single method per allergen ingredient leaves testing workflows vulnerable to not detecting the presence of an allergenic ingredient due to the differing limitations of all methods. Workflows have therefore been developed to support incident management, combining different technology formats in order to maximise the scope of the testing for each allergen and applicability depending on the level of processing (Section 9). To increase our understanding of allergen risk management in the supply chain, stakeholder interviews are included in this report detailing the various strategies used across UK suppliers to manage food allergen risk.

10.4 Future direction

In Japan, much progress has been made to align food regulation with allergen testing methods. While this approach may not be perfect, especially since the range of permitted test methods is mandated, thus perhaps stifling innovation while also removing some commercial testing methods from the market, this may be a useful model aiming to standardise testing and offer improved protection to consumers.

Since a different model is used compared to that in the UK and EU, there may be areas of learning that can be made from discussing allergen management with countries which have adopted a contrasting approach. For example, it is possible that establishing an approval system where manufacturers submit their validation data to be assessed by experts, or that a requirement is made for companies to publish up-to-date validation and performance data with their kits, would bring benefits to UK procedures for food testing.

The Japanese are recognised as being advanced compared to other nations regarding management of allergens and also have different views regarding how RMs should be prepared (personal comms) or how calibrants should be applied and it may be interesting to learn more on their methods in an interview with relevant academics in Japan.

As highlighted by many stakeholders and our expert consultants, there is a need for the development and commercialisation of a fast, accurate, low-cost, multi-allergen tests, which can be used at point-of-use in factories to identify allergens. One candidate rapid technology may be Near Infrared Spectroscopy (NIR) although development of such methods is required.

While ELISA, lateral flow test strips and PCR testing methods have dominated the allergen testing scene for a number of decades and therefore benefit from years of method development and validation, new methods are becoming increasingly apparent, with increased specificity and robustness to thermal processing. Mass spectrometry methods fall under this category. While LC-MS methods show benefits in specificity compared to ELISA, sensitivity is often lacking in these relatively new mass spectrometry methods at present. However, some mass spectrometry methods deliver to the expected guidelines of FAO/WHO and, as highlighted in the EFSA ThRAll project, method optimisation for certain allergens has the potential to further improve sensitivity. In order to progress more recent methods of allergen detection, increased funding for R&D is required.

Concerning emerging technologies, methods using alternative technologies are under development but are not being disclosed in the public domain at present for proprietary reasons(personal comms). It would be beneficial to discuss these technologies in confidence, to better inform FSA of emerging technologies. No matter the platform (e.g. ELISA, PCR, lateral flow, LC-MS/MS, emerging technologies), the development of reliable, rapid and point-of-use allergen detection methods showing high accuracy and sensitivity is paramount to protect allergen-sensitive consumers.