Is food allergen analysis flawed? Health and supply chain risks and a proposed framework to address urgent analytical needs

M. J. Walker,*a D. T. Burns,b C. T. Elliott,b M. H. Gowlandc and E. N. Clare Millsd

Food allergy is an increasing problem for those affected, their families or carers, the food industry and for regulators. The food supply chain is highly vulnerable to fraud involving food allergens, risking fatalities and severe reputational damage to the food industry. Many facets are being pursued to ameliorate the difficulties including better food labelling and the concept of thresholds of elicitation of allergy symptoms as risk management tools. These efforts depend to a high degree on the ability reliably to detect and quantify food allergens; yet all current analytical approaches exhibit severe deficiencies that jeopardise accurate results being produced particularly in terms of the risks of false positive and false negative reporting. If we fail to realise the promise of current risk assessment and risk management of food allergens through lack of the ability to measure food allergens reproducibly and with traceability to an international unit of measurement, the analytical community will have failed a significant societal challenge.

Three distinct but interrelated areas of analytical work are urgently needed to address the substantial gaps identified: (a) a coordinated international programme for the production of properly characterised clinically relevant reference materials and calibrants for food allergen analysis; (b) an international programme to widen the scope of proteomics and genomics bioinformatics for the genera containing the major allergens to address problems in ELISA, MS and DNA methods; (c) the initiation of a coordinated international programme leading to reference methods for allergen proteins that provide results traceable to the SI.

This article describes in more detail food allergy, the risks of inapplicable or flawed allergen analyses and propose a framework to address the currently unmet and urgently required analytical requirements. Support for the above recommendations from food authorities, business organisations and National Measurement Institutes is important; however transparent international coordination is essential. Thus our recommendations are primarily addressed to the European Commission, the Health and Food Safety Directorate, DG Santé. A global multidisciplinary consortium is required to provide a curated suite of data including genomic and proteomic data on key allergenic food sources, made publically available on line.

Introduction

Food fraud poses risks to consumer confidence and business continuity, as the horse meat episode demonstrated,1 and also to health, for example the well recorded morbidity and mortality from counterfeit alcoholic drinks.2–4 There are also risks to people with food allergies as we will describe. Allergen analysis is a key tool in supply chain scrutiny as well as in the investigation of adverse reactions. Herein we briefly introduce food allergy, describe with examples the risks of inapplicable allergen analyses and propose a framework to address the currently unmet and urgently required analytical requirements.

Food allergy

Food allergies, i.e. adverse immunologic (IgE and non-IgE mediated) reactions to food, have resulted in considerable morbidity5 and reached epidemic proportions in the industrialized world6,7 affecting up to 10% of young children and 2–3% of adults. Anaphylaxis, a rapid onset multi-organ system allergic reaction with release of chemical mediators from mast cells and basophils, can cause fatalities. The risk of such deaths, though comparatively rare,8 contributes to well-docu-
ment detriment to the quality of life for allergic consumers and their families.9–11 There are burdens on health care,12 on businesses (food recalls, for example) and regulators13 and in less developed countries where, owing to poor labelling and awareness, significant challenges may exist. Current reputed cures for food allergies remain experimental and lifelong avoidance of the eliciting food(s) is required. Food intolerance such as coeliac disease also imposes significant burdens14 and strict food avoidance is usually necessary.

Legislation, risk and thresholds

Regulatory risk management strategies for allergic consumers have focused on providing information about the presence of food allergens through label declarations.15 In December 2014 European labelling law, Regulation 1169/2011, extended such disclosure requirements, see Table 1, to non-prepacked food including that available in catering establishments.16 Cross-contamination with allergens may trigger the general principles of European17 and UK food law (Food Safety Act 1990) that make it an offense to sell food that is unsafe for, or not of the nature, substance, or quality demanded by allergic consumers, particularly if specifically intended for their consumption. The risks posed by the unintended presence of allergens in food have resulted in proliferation of precautionary (e.g. “may contain...”) labelling, widely regarded as unsatisfactory.18,19 Much effort is being expended on the development of ‘thresholds’,20 ‘action levels’,21 or ‘reference doses’, e.g. those advocated by the European Academy of Allergy and Clinical Immunology, EAACI22 for major allergens. The lack of agreed limits below which only the most sensitive allergic subjects might react is viewed by many as preventing the development of evidence-based allergen management strategies that are understood by clinicians, patients and industry. The European Food Safety Authority (EFSA) has comprehensively reviewed allergenic foods and food ingredients for labelling purposes. EFSA however declined to consider advocating thresholds on the grounds that labelling and the level of risk that may be acceptable (e.g. the fraction of the allergic population protected and to what extent) are risk management decisions outside EFSA’s remit.23 Nevertheless, the principal legal measure16 governing food labelling in the EU includes powers (Article 36) for the European Commission to adopt law on information on the unintentional presence in food of substances causing allergic or intolerance based reactions. The most probable basis for such law would be allergen elicitation thresholds which are not possible without allergen reference materials. Table 2 shows reference doses cited by the European Academy of Allergy and Clinical Immunology, EAACI, these data (second column) are identical to the data exhibited in the Allergen Bureau Voluntary Incidental Trace Allergen Labelling, VITAL® scheme24 except for fish.

Supply chain vulnerability – allergens

Two incidents illustrate well the vulnerability of the food supply chain, and hence of allergic consumers. These examples (a) illustrate evidence of deliberate substitution of almond by peanut in the supply chain and (b) describe what was initially thought to be deliberate adulteration of cumin with almond but

As Managing Director of Michael Walker Consulting Ltd Michael has a portfolio of roles. He is Consultant Referee Analyst in the Laboratory of the Government Chemist, where he also manages research on allergen measurement and facilitates a Food Allergy & Food Intolerance Knowledge Network. He was a subject matter expert to the UK DH/Defra Elliott Review into the integrity and assurance of food supply networks. Michael is member of the European Academy of Allergy & Clinical Immunology and Chair of the FSA NI Strategic Committee on Food Surveillance. Michael also has a thriving chemico-legal practice and extensive non-executive director experience including as a founder board member of the FSA. Michael’s background is in chemistry. He is a Fellow of both IFST and RSC, and holds the MChemA, the statutory qualification in food drugs and water to act as a Public Analyst in the UK.

Professor Duncan Thorburn Burns MRIA FRSE, became Emeritus Professor of Analytical Chemistry in The Queen’s University of Belfast in 1999. He remains internationally professionally and academically active with 100 peer reviewed papers and 3 books since formal retirement, and regularly appears as an expert witness. Currently he is an Honorary Research Professor, resident in the Institute for Global Food Security at QUB. His on-going interests are in the interpretation and validity of analytical data for legal/forensic purposes which ties in with his involvement in the research and publication carried out in the Government Chemist Programme of LGC.
in fact turned out to be contamination of the cumin supply chain with mahaleb. Both examples demonstrate that good allergen analysis is necessary to help protect the supply chain.

Almond or peanut?

Enforcement surveillance\(^2^4\) of allergen compliance in catering establishments regularly concludes that specifically asking for an allergen-free meal provides little real protection. Such appeared to be the case for two chicken tikka masala meals found to contain peanut in a survey in 2010/11.\(^2^5\) However, follow up revealed a supplier had introduced groundnuts (peanuts) instead of almond powder as contracted into the supply chain. The firm was convicted on prosecution. The conviction was, however, overturned on appeal on technical legal grounds.\(^2^6\) Nevertheless the Food Standards Agency (FSA) Annual Report of Incidents 2012 refers to investigations of severe allergic reactions following the consumption of curry dishes purchased from Indian restaurants and takeaways. Noting that some of these incidents resulted in fatalities, FSA reported that some incidents were caused by the use of a ground almond ingredient, which also contained ground peanut (groundnut). FSA identified weaknesses in the food chain where such contamination and loss of clear information occurred, including poor understanding of the significance of substituting peanuts for almonds, incorrect allergen information provided at a point of sale, and unclear labelling and confusion between peanuts and tree nuts (almonds) leading to the potential for accidental substitution. However FSA also reported possible economically motivated adulteration, driven by the financial incentive to substitute ground almonds with ground peanut.\(^2^7\)

Almond or mahaleb?

Against this backdrop in October 2014 when Canadian authorities found undeclared peanut and almond protein in products containing cumin, it was feared that a further, potentially life threatening, breach of supply chain security had occurred.\(^2^8\) Suspicions fell on cross contamination in harvesting, transport, storage, or processing, adulteration with almond shells, or in the case of the peanut protein, adulteration with peanut hulls or peanut meal animal feed. The FSA issued the first of a small number of related recalls, of ground cumin sold by the Barts Ingredients Company Ltd to contain traces of almond protein not listed on the label, on 31 January 2015.\(^2^9\) FSA referred this as an official technical appeal\(^3^0\) to the Government Chemist,\(^3^1\) asking for a review of the analysis that had led to the recall. In early March 2015 Barts Ingredients Company Ltd claimed publicaly that...
another material, mahaleb, gives a positive reading for almond using test methods. On 30 April 2015 the Canadian authorities rescinded product recalls of cumin and cumin-containing products previously thought to contain undeclared almond. The Canadian statement noted that the recalls had been based on “original laboratory results [that] were false positives...[caused by] cross-reactivity of mahaleb...(Prunus mahaleb), with the almond allergen test kit. It is highly likely that the positive sample results for the ground cumin and cumin-containing products were due to mahaleb contamination and not almond”. Almond is a member of the genus ‘Prunus’ – trees and shrubs, which includes plums, cherries, peaches, nectarines, apricots and mahaleb. Prunus mahaleb was previously little known in the UK but was said also to have been handled in the cumin supply chain. The UK Government Chemist subsequently determined that although limitations still remain in the state of the science that prevent the presence of almond being completely ruled out, the results of the technical appeal investigation indicate that the queried sample contained a Prunus protein and DNA the origin of which was consistent with mahaleb rather than almond.

While P. mahaleb does not appear in the list of allergens required by law to be declared if used intentionally as an ingredient in food, it is important for a food business to understand its supply chain to assess and manage cross contamination risks. Given the amino acid homology between almond and mahaleb a risk to Prunus-allergic individuals might remain.

Prior to taking up a chair in Molecular Allergology at the University of Manchester in September 2011, Clare led the BBSRC Institute Strategic Programme Grant on Food Structure and Health and was co-ordinator of the EU project EuroPrevall. After moving to Manchester she led the successful bid to the EU for the iFAAM project, now in its third year. Her personal research interests are focused on structure–function relationships in food proteins particularly what makes some proteins, and not others, become allergens, including the effects of the food matrix and processing on resistance of food proteins to digestion and the role this plays in determining the allergenicity of foods. Clare serves on the Advisory Committee on Novel Foods and Processes, is Chair of the Food Allergy Interest Group Board, European Academy of Allergy and Clinical Immunology, is a Fellow of the International Academy of Food Science and Technology of The International Union of Food Science and Technology and a member of the Self-tasking Allergenicity working group of the EFSA GMO panel. After a PhD from the University of Kent at Canterbury Clare held research posts in industry, the Department of Health and the Institute of Food Research, Norwich with an honorary Chair, Dept Pharmacy, University of East Anglia (until 2011) and is currently a visiting Professor in Wageningen University.

### Table 1 Allergens covered by EU labelling law, Annex II Regulation 1169/2011

<table>
<thead>
<tr>
<th>Annex II entry</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cereals containing gluten and products thereof</td>
<td>Wheat, Rye, Barley, Oats, Shrimp/prawn, Crab, Lobster, Crayfish</td>
</tr>
<tr>
<td>Crustaceans and products thereof</td>
<td>Skinned milk powder, Cheese etc.</td>
</tr>
<tr>
<td>Fish and products thereof</td>
<td>Almond, Hazelnut, Walnut, Cashew, Pecan, Brazil, Pistachio, Macadamia</td>
</tr>
<tr>
<td>Peanuts and products thereof</td>
<td>Celery and products thereof, Mustard and products thereof, Sesame seed and products thereof, Sulfur dioxide/sulfites, Lupin and products thereof, Molluscs and products thereof</td>
</tr>
<tr>
<td>Milk and products thereof</td>
<td>Mussels, Scallops, Cockles, Oyster, Clam</td>
</tr>
</tbody>
</table>

There are limited exceptions cited in Annex II of ingredients that do not contain sufficient allergenic protein to elicit a reaction.

### Table 2 EAACI reference doses and suggested clinically relevant reference material concentrations

<table>
<thead>
<tr>
<th>Food</th>
<th>EAACI reference dose(^a)</th>
<th>Suggested clinically relevant(^b) RM allergen protein concentrations mg kg(^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peanut ED 1%</td>
<td>0.2 mg peanut protein</td>
<td>2–10</td>
</tr>
<tr>
<td>Cow’s milk ED 1%</td>
<td>0.1 mg milk protein</td>
<td>1–10</td>
</tr>
<tr>
<td>Egg ED 1%</td>
<td>0.03 mg egg protein</td>
<td>0.3–5</td>
</tr>
<tr>
<td>Hazelnut ED 1%</td>
<td>0.1 mg hazelnut protein</td>
<td>1–10</td>
</tr>
<tr>
<td>Soya ED 5%</td>
<td>1.0 mg soya protein</td>
<td>10–100</td>
</tr>
<tr>
<td>Wheat ED 5%</td>
<td>1.0 mg wheat protein</td>
<td>10–100</td>
</tr>
<tr>
<td>Cashew ED 5%</td>
<td>2.0 mg cashew protein</td>
<td>20–100</td>
</tr>
<tr>
<td>Mustard ED 5%</td>
<td>0.05 mg mustard protein</td>
<td>0.5–5</td>
</tr>
<tr>
<td>Lupin ED 5%</td>
<td>4.0 mg lupin protein</td>
<td>40–200</td>
</tr>
<tr>
<td>Sesame seed ED 5%</td>
<td>0.2 mg sesame protein</td>
<td>2–10</td>
</tr>
<tr>
<td>Shrimp ED 5%</td>
<td>10 mg shrimp protein</td>
<td>100–1000</td>
</tr>
<tr>
<td>Fish ED 5%</td>
<td>0.1 mg fish protein(^a)</td>
<td>1–10</td>
</tr>
</tbody>
</table>

ED \(x\) %, eliciting dose for \(x\) % of the allergic population. \(^a\) Provisional. \(^b\) Assuming a minimum portion size of 100 g.

**E. N. Clare Mills**
Allergen analysis

Analysis for food allergens is required for many reasons. Key industry standards emphasise greater transparency, traceability and integrity in the supply chain requiring analysis to check that food is what it is claimed to be, and encourage systems to reduce exposure to fraud. Analysis supports validation and verification of factory cleaning and investigation of recalls and incidents. Surveillance and enforcement, particularly after the introduction of more extensive labelling requirements, rely heavily on analysis to support and protect consumers and responsible businesses and, in the event of adulteration, provide evidence for criminal or civil action in the courts; a key deterrent. Investigation of adverse reactions may require analysis to find out what caused the reaction, and therefore enable the individual to avoid it in the future. Investigation of fatalities, already problematic, requires analysis e.g. of food seized at the incident, stomach contents or other forensic exhibits.

The recindment, (see above) because of initially flawed analysis, of over two dozen allergen recalls on both sides of the Atlantic risks uncertainty and confusion over allergen testing in the future. This jeopardises consumer safety now and the development of allergen thresholds in the future. The origins and resolution of these problems lie in the difficulties of allergen analysis.

Food allergens that bind to IgE are large protein molecules and many approaches have been taken to their analysis. Most routine food allergen analysis is undertaken by Enzyme Linked Immunosorbent Assay (ELISA) enabling detection and (semi-) quantification. Polymerase Chain Reaction (PCR) assays are also applied in allergen risk assessment and management. For both techniques detection is less of an issue, although not without problems but sound quantification remains elusive. Commercially available ELISA kits exhibit variable and manufacturer specific sensitivities and cross-reactivity. In proficiency testing multimodal datasets for allergen ELISAs are common and different assigned values have to be generated for the different kits used. Recent work on precautionary labelling on pre-packed processed food and concentrations of certain cross contaminant allergens in foods suffered from unexpected cross reactivity in the commercial ELISA. False positive results were identified arising from an Association of Analytical Communities (AOAC) approved peanut assay owing to cross reactivity to soya. The cross reactivity which was evidently not a feature of the original assay seemed to have developed after many years use of the kit and necessitated a troublesome late stage review of the research findings. Structural changes in the target molecules by food processing or sample extraction may prevent detection. PCR assays are probative of the source species DNA (which may not be present e.g. egg white) rather than the allergen protein. Moreover proteins are the hazard and thus the key measurand. PCR is essentially qualitative at present. Quantification based on copy number can be derived from cycle thresholds but requires reference materials to construct a calibration curve, although digital PCR may circumvent this difficulty. Even so is not easy to convert a quantification based on copy number to a weight/weight basis. There has been little systematic research on the relationships between the findings of PCR approaches and protein techniques such as ELISA or liquid chromatography and tandem mass spectrometry, LC-MS/MS, for allergen analysis to assess their comparability. Hence, there is a requirement for orthogonal methods that confirm molecular identity and that are capable of valid quantification. LC-MS/MS methods, e.g. multiple reaction monitoring of peptides arising from enzymatic digestion of proteins, offer such advantages, along with the possibility of multiplexed high throughput. The application of LC-MS/MS is still recent in food allergen analysis. It is possible to detect proteins and peptides with a high degree of sensitivity and resolving power, providing protein composition, structure and sequence information, and MS has the potential for a wide linear dynamic range, and absolute identification and quantification of allergens. However the techniques require a high level of expertise and costly equipment; extraction and cleanup steps are necessary and the methods can be laborious and time consuming. The complexity of most food matrices represents a significant challenge even to MS although guidance is available, including on a model system that demonstrates isotope dilution mass spectrometric traceability from a set of peptides to an allergenic protein. Because ELISA is much more widely used for allergen analysis than MS or PCR there is more published evidence of its deficiencies but similar deficiencies apply to both MS and PCR approaches. Thus the promise of MS or PCR will be lost if underpinning work suggested herein is not carried out.

In summary, current allergen analysis would be impossible without ELISA which has brought many benefits in allergen risk assessment and risk management. However all current forms of allergen analysis present some deficiencies which may jeopardise present and future risk assessment and risk management of food allergy, a problem of high and increasing importance.

The attitude of the analytical community

In order to assess the attitude of the analytical community in the UK to food allergen analysis a small anonymous survey of prominent food analytical service providers, including Public Analysts was carried out (8/07/2015–14/07/2015). The responses (n = 36) showed that food allergen analysis is very (88.9%) or somewhat (11.1%) important to laboratories, and is increasing (63.9%) or increasing a lot (11.1%). Carrying out such analyses was deemed to be difficult (41.7%) or extremely difficult (2.8%) and 48.6% found reference materials (RMs) to be available but with reservations as to their usefulness or credibility, available but of limited relevance (22.9%) or not available (17.1%). Confirmation of food allergen detection and quantification by an orthogonal technique was scored as very difficult (45.7%) or slightly difficult (25.7).
A framework to address the problems

The importance of allergen analysis, its relative difficulty and lack of relevant and trusted reference materials revealed in the above survey correspond with the experiences of the authors. Three distinct but interrelated areas of analytical work are urgently needed to address the substantial gaps identified; production of reference materials, a bioinformatics gap analysis and development of reference methods resulting in metrologically traceable results.

Reference materials

Sykes et al.44 showed that inclusion of a ‘reference spiked sample’ in a Proficiency Test, PT, round where the raw data were non-normal and multi-modal, tended to yield ratio data that were normal and symmetrically distributed. These authors, and (non-exhaustively) many others51–54 call for the development of internationally recognised sets of allergen reference materials to improve the reliability of allergen analysis. Reference material (RM) and Certified Reference Material (CRM) are well defined terms55–57 within an associated international infrastructure.58 It is not always clear that the limited number of food allergen RMs currently commercially available comply with this infrastructure. Reference materials produced by National Measurement Institutes exhibit the highest standards. Taylor et al.59 described with examples the preparation of naturally incurred standards as allergenic food residues incorporated into various representative food matrices and then processed in a manner similar to ‘real-world’ food processing.

Making a reference material is relatively expensive owing to the complexity of production. The following steps should be carried out within a documented quality system:

- Effective project planning and project management
- Definition of need, background and clear specification
- Material procurement
- Identification or development of a validated analytical method for the measurand so as to distinguish measurement dispersion from dispersion arising from homogeneity and stability issues
- Preparation of the material by well characterised methods
- Packaging to ensure integrity and stability
- Storage under controlled conditions to maintain stability
- Homogeneity and stability evaluation by validated methods with known performance data
- Characterisation and certification (if that proves possible) with documented traceability of values, an uncertainty budget and consideration of the commutability of the material
- Preparation of a certificate to accompany the material, and a production report
- Distribution and sales, ensuring integrity of the material
- On-going monitoring and customer support.

Thus producers of allergen RMs should address the above points and attempt to ensure the matrix is industrially realistic for processed food. The incurred concentrations should be appropriate for and preferably establish a relationship with the concentrations that affect allergy sufferers. A prototype such material set, (a blank material and a QC material with peanut protein added at 10 mg kg$^{-1}$) has been prepared based on a EuroPrevall study matrix (chocolate dessert mix)60 used to assess clinical thresholds.61,62

Both calibrant and matrix reference materials for food allergens are required. However their production is not trivial. The legislation defines allergens in terms of the food (Table 1) but analysis targets proteins or their peptides (or DNA). For the protein allergens the analyte is often neither exactly defined nor easy to render identical in sample and calibrator. Typically, multiple allergen proteins and isoforms are present, in a complex matrix. Taking peanut as an example, the food itself includes proteins, lipids, carbohydrates and minerals available in multiple processed formats including raw, roasted and/or defatted to varying degrees and included in a wide range of other foods.64 The peanut allergens include at least 12–14 multiple specific proteins, of which only Ara h 1, Ara h 2, Ara h 3, and Ara h 6 have been demonstrated to be clinically important.65,66 Protein post-translational modifications, PTM, occur and further complexity is introduced by biological variation, fractionation (intended and adventitious) and reaction with other food components. The analyte (measurand) therefore may be, in MS, peptides expected to be uniquely representative of specific proteins, for ELISA, known proteins that may or may not be the allergens or, for PCR, a DNA sequence. Pragmatism is required52 as ideal solutions to the above problems will not easily, economically or soon be found. Therefore a staged approach is needed starting with non-ideal reference materials, as explained below. Maximum transparency is required as to the commercial origin and compositional characteristics of the allergenic food used to formulate, initially, simple matrix reference materials gravimetrically prepared at blank (zero) and clinically relevant allergen concentrations. Such concentrations are suggested in Table 2. Homogeneity and stability studies and further characterisation by, at least, ELISA should be performed. Experience gained will enable progression to incurred allergens in processed foods representing a suitable spectrum of protein, lipid and carbohydrate compositions,67 followed by production of certified reference materials representing those RMs found most useful. It should be noted that production of reference materials is rarely a commercial proposition.

Much can be learned from work that led to what remains the ‘gold standard’ reference material for gliadin, described in the proceedings of the Working Group on Prolamin Analysis and Toxicity (WGPAT).68 Securing food that is free from gluten (gliadin) for those with coeliac disease is as important as an allergen-free diet and fraught with the same analytical difficulties. Moreover the definition of gluten is empirical: “gluten” is defined as a protein fraction from wheat, rye, barley, oats or their crossbred varieties and derivatives, to which some persons are intolerant and that is insoluble in water and 0.5 M NaCl. The prolamin content of gluten is generally taken as 50% and prolamins are defined as the fraction from gluten
that can be extracted by 40–70% ethanol.\(^6^9\) This definition enabled WGPAT to prepare a gliadin reference material by extraction from milled wheat kernels representing a specific year’s harvest of the most commonly grown cultivars in 3 European countries. Moreover the obtained gliadin (PWG-gliadin) was characterised by a wide spectrum of techniques including immunological, MS and electrophoretic as well as for stability and homogeneity, as summarised by van Eckert \textit{et al}.\(^7^0\) In 2005 the Institute for Reference Material and Measurements of the European Commission (IRM\texttext{M}) declined to accept PWG-gliadin as a certified reference material and returned it to WGPAT in 2006 from where it can be obtained.\(^6^8\) Although it is difficult to speculate on the reasons for IRMM’s action, the want of a route to full metrological traceability for PWG-gliadin, (and allergens in general), that the proposals in this paper seek to address, may be one of the root causes.

It is recommended that a coordinated international programme be set up for the production of properly characterised reference materials and calibrants for allergen analysis, beginning with the rapid availability of simple materials \textit{e.g.} containing, separately, the major allergens (\textit{e.g.} as defatted and/or freeze dried powdered substances) (Table 1) at zero (blank) and clinically relevant concentrations (Table 2). The programme should progress to incurred allergens in processed foods representing a suitable spectrum of protein, lipid and carbohydrate compositions,\(^6^7\) followed by production of certified reference materials representing those RMS found most useful.

\textbf{Bioinformatics gap analysis}

Prior to the Government Chemist’s investigation that led to the recent rescindment of the UK recall of cumin, cross reactivity of some commercial almond ELISAs to apricot kernel was well known and acknowledged.\(^7^1\) That such cross reactivity was widespread across the genus Prunus only became apparent during our investigation, and resulted from the homologies across that genus. DNA databases of NCBI\(^7^2\), GenBank,\(^7^3\) and BOLD\(^7^4\) were accessed for publically available DNA and amino acid sequence data on the genus Prunus. Searches (24/03/15) revealed 904 individual records of Prunus on the BOLD database, representing 188 different species originating from 19 different countries. A search of the GenBank database revealed 236,565 nucleotide sequences for Prunus, of which 99,491 are derived from \textit{P. armeniaca}, 1747 from \textit{P. dulcis} but only 49 from \textit{P. mahaleb}.\(^3^6\) It was also required to supplement information from \textit{e.g.} UniProt\(^7^5\) which provides freely accessible protein sequence and functional information that is as comprehensive as currently possible, with experimental data on suitable peptides by LC-Time-of-Flight MS. Hence there is a gap in the bioinformatics databases that places difficulties in the way of designing PCR DNA and MS assays with the required specificity to manage the supply chain for allergens. To our knowledge no systematic assessment is available of such bioinformatic gaps and homology across the species for the priority major allergens.

It is recommended that an international programme be set up to widen the scope of nucleotide and amino acid sequence data-bases and proteomics and genomics bioinformatics generally for the genera containing the major allergens (Table 1) to pinpoint possible future problems in ELISA, MS and DNA methods. Further coordinated international programmes to mitigate any cross reactivity and fill database gaps are also recommended.

This task can be made easier by emerging genomic and proteomic tools that foster better understanding of molecular and phylogenetic relationships and conservation of sequences. It is possible to impute missing data on nucleotide base sequences and protein amino acid sequences in an efficient and effective manner. Moreover, allergen bioinformatics would benefit from more efficient search tools such as text mining and ‘Intelligent Systems for Molecular Biology’.\(^7^6\) Learned journals also have a part to play by requiring uploading to databases mentioned above and, for example PRIDE,\(^7^7\) of relevant information (\textit{e.g.} mass spectrometry proteomics data) as a condition of publication of a submitted manuscript.

\textbf{Metrological traceable methods}

Metrological traceability is the property of an analytical result that allows measurements made in different laboratories under different conditions to be compared in a meaningful way, within an international infrastructure, the International System of Units (the SI). Such work is carried out by National Measurement Institutes (NMs), in each developed country.\(^7^8\)–\(^8^0\) Metrological traceability of allergen protein data is currently possible only by MS-based absolute quantification such as isotope dilution MS, IDMS, a primary ratio method that relates results directly to the SI with a small measurement uncertainty, and which is commonly used for the characterisation of small molecule CRMs.\(^8^1\) The principles of exact matching (EM)-IDMS have been applied to absolute quantification of proteins based on proteolytic (most commonly tryptic) digestion of the protein, the use of isotopically labelled peptides as internal standards and of synthetic unlabelled peptides as primary standards. Isotopically labelled and unlabelled peptides are more readily available, and less expensive than isotopically labelled proteins, and are better characterised. Application to allergen proteins is difficult and costly,\(^5^0\) but once achieved can be cascaded \textit{via} reference materials and certified reference materials so that the outcomes should be available at modest cost to support routine analysis.

It is recommended that an international programme be initiated leading to reference measurement methods for allergen proteins which provide results traceable to the SI. The methods should be applicable to the major allergens (Table 1), at clinically relevant concentrations (Table 2) in processed foods covering an appropriate range of protein, lipid and carbohydrate composition.

\textbf{International collaboration}

National support from food authorities, business organisations and National Measurement Institutes for the above recommendations is important, however international coordi-
nation is essential. The European Union legislates for the largest number of priority food allergen groups (Table 1).\textsuperscript{15} Within the European Commission, the Health and Food Safety Directorate, DG Santé, is responsible for protection and improvement of public health, ensuring Europe's food is safe and wholesome and that citizens can be confident that their interests are protected. Work is already underway by several bodies, including individual National Measurement Institutes, (including LGC, the Joint Research Centre (JRC) of the European Commission and NIST), the MoniQA Association\textsuperscript{82,83} and iFAAM, a European Union Seventh Framework Programme ‘Integrated approaches to food allergen and allergy management’.\textsuperscript{84} Collating the various global work streams is needed to focus on the interrelated areas of analytical work that we have identified. This must be done in a transparent manner to achieve the aspirations of all stakeholders. DG Santé fulfils the criteria suggested by the above analysis. Thus our recommendations are primarily addressed to DG Santé which should work closely with relevant bodies outside Europe to avoid duplication of effort or gaps.

**Conclusions**

Food allergy is an increasing problem for all stakeholders and the food supply chain has been shown to be vulnerable to fraud involving food allergens, risking fatalities and reputational damage to the food industry. Legislation, risk assessment and risk management of food allergens show a high dependency on the ability to detect food allergens and quantitatively determine them. All current analytical approaches exhibit described deficiencies that jeopardise accurate results and risk false positives and false negatives. If we fail to realise the promise of many strands of risk assessment and risk management of food allergens through lack of the ability to measure food allergens reproducibly and with traceability, the analytical community will have failed a significant societal challenge. We recommend three distinct but interrelated areas of work urgently needed to address the substantial gaps identified: reference material production and reference method development along with better bioinformatics. There are multiple strands of risk assessment and risk management currently underway however the participant organisations should increase their mutual interaction. Moreover the scale of the problems identified and their technical solution are such that only a planned international programme, coordinated we suggest by DG Santé will be capable of addressing the issues quickly and efficiently to provide integrated solutions. A global multidisciplinary consortium is required to provide a curated suite of data including genomic and proteomic data on key allergenic food sources, made publically available on line. As a first step we suggest a pump priming workshop within the next year to support the development of such a consortium bringing together the analytical, food science and clinical communities along with patient support groups, representatives of food manufacturers, and the regulatory agencies. Such an international programme should be aligned with the timescale envisaged for the implementation of Article 36 of Regulation 1169/2011 that gives the Commission powers to address allergen cross contamination. We suggest EFSA, following its opinion on allergens,\textsuperscript{23} should continue to be involved; the programme should be of at least 5 years duration and

![Fig. 1](image_url) Simplified diagrammatic ‘traffic light’ illustration of work required. (PTM – Post translational modification).
must have regard to the work of the organisations and research programmes mentioned in this paper but also e.g. the Allergen Bureau,21 and EAACI22 in the development of reference doses.

It is clear that there are significant problems to be solved, for example do we know if proteins purchased as a starting step in an analytical investigation really mimic the allergenic proteins e.g. as regards post translational modification, PTM, and tertiary structure? However, with work on these and all the strands outlined herein progress can be made. Calibrants are needed such as gravimetrically prepared peptide solutions with known concentrations traceable to the SI, or a solution of a well characterised protein of known concentration traceable by way of peptides to the SI.49,50 But how will these relate to a matrix reference material, say a food such as light roasted defatted peanut incurred in an industrially relevant matrix at a clinically relevant concentration? Fig. 1 illustrates in a highly simplified manner how this might be accomplished. For some food allergens clinical and bioinformatics studies have already identified relevant markers or allergenic proteins for which signature peptides are available. But this remains to be accomplished for all the major allergens. With the identification of the major relevant proteins of an allergenic food, and characterisation of the impact of food processing, analytical extraction, PTM, and tertiary structure (none of these are trivial tasks) a reference material can be created by either of two related approaches:

(a) A ‘chimera’ theoretical matrix RM containing, in an industrially relevant matrix, clinically relevant concentrations of the optimal number of the separate component allergenic proteins of the food allergen already individually characterised and traceable to the SI by isotope dilution MS of the signature peptides, Cryar et al.50 following the recommendations of Johnson et al.49 or

(b) An empirical matrix RM containing, in an industrially relevant matrix, clinically relevant concentrations of mixed proteins extracted and characterised as described by van Eckert et al.75 and further traceable to the SI by investigations described in (a).

Fig. 1 also indicates by way of a ‘traffic light’ code current progress towards the above goals; ‘green’ (i.e. accomplished), ‘amber (i.e. under way) or ‘red’ (i.e. yet to be done).

The recommendations herein are of a complexity and resource demand that only an internationally coordinated effort can accomplish them. However, rarely has such an exciting interdisciplinary scientific endeavour arisen as a solution to a key socially relevant problem.

Acknowledgements

Helpful discussions with LGC colleagues Dr Milena Quaglia, Science Leader, Protein Mass Spectrometry, Chief Chemical Metrologist Dr Mike Sargent and Head of Reference Material Production Mrs Gill Holcombe are gratefully acknowledged. The authors are very grateful to the reviewers whose critical appraisal and helpful suggestions have improved the quality of this paper.

The work described in this paper was part funded (MJW) by the UK Department of Business Innovation and Skills as part of the Government Chemist Programme 2014–2017 however the views expressed in this paper are those of the authors alone. The paper should not be taken as an authoritative statement or interpretation of the law, as this is a matter for the courts.

References


68 Working Group on Prolamin Analysis and Toxicity (WGPAT), http://www.wgpat.com/aims.html, accessed 19.10.2015 (The earlier proceedings may be available in paper copy from the WGPAT with electronic copies from 2010 onwards however an excellent summary is in reference 70 below).


71 See for example kit instructions RIDASCREEN®FAST Mandel/Almond R690, R-Biopharm AG, Landwehrstrasse 54, D-64293 Darmstadt, Germany.

72 NCBI: National Center for Biotechnology Information.

73 GenBank®: US National Institutes of Health genetic sequence database, an annotated collection of all publicly available DNA sequences. It is part of the International Nucleotide Sequence Database Collaboration, which comprises the DNA Database of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI.

74 BOLD: Barcode of Life Data System.


76 See for example the proceedings of the joint 23rd annual meeting of Intelligent Systems for Molecular Biology (ISMB) and 14th European Conference on Computational Biology (ECCB), ISMB/ECCB 2015, July 10 to July 14, 2015, Dublin, Ireland in Bioinformatics, 31 (12), 2015.


84 University of Manchester, Integrated approaches to food allergen and allergy management (iFAAM) http://www.inflammation-repair.manchester.ac.uk/iFAAM/, (accessed 15.07.2015).