Classification and Coding of Data About IgE-mediated Food Allergic Reactions

Chloe French

The University of Manchester Manchester Institute of Biotechnology

Benjamin Green

The University of Manchester Manchester Institute of Biotechnology

Saskia Lawson-Tovey

The University of Manchester

Bushra Javed

The University of Manchester Manchester Institute of Biotechnology

Martine Morisset

University Hospital Centre Nancy: Centre hospitalier regional universitaire de Nancy

Andre Knulst

University Medical Centre: Universitair Medisch Centrum Utrecht

Thuy-my Li

University Medical Centre: Universitair Medisch Centrum Utrecht

Angela Simpson

The University of Manchester Faculty of Biology Medicine and Health

Shelly Dua

Cambridge University: University of Cambridge

Nicolette de Jong

Erasmus Medical Center: Erasmus MC

Barbara Ballmer-Weber

University Hospital Zurich: UniversitätsSpital Zurich

Montserrat Fernandez-Rivas

Hospital Clinico Universitario San Carlos

Kirsten Beyer

Charite Universitätsmedizin Berlin Campus Charite Mitte: Charite Universitätsmedizin Berlin

Clare Mills (✉ clare.mills@manchester.ac.uk)

Manchester Institute of Biotechnology, School of Biological Sciences, Division of Infection, Immunity and Respiratory Medicine, Manchester Academic Health Science Centre, University of Manchester, UK.

Allergy <https://orcid.org/0000-0001-7433-1740>

Philip Couch

The University of Manchester

Research

Keywords: Ontology, Terminology, Vocabulary standards, Data harmonisation, Food Allergy, Hypersensitivity, ThRAI

Posted Date: December 7th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-118179/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Background: Collation of clinical data on IgE-mediated food allergies is essential to provide evidenced-based approaches to managing and treating food allergies and prevent accidental reactions. However, this can be a time consuming and difficult process due to the heterogeneous way in which studies collect such data. In order to facilitate data harmonisation a set of standardised terminologies have been identified and a consensus technique established to code food allergy data.

Methods: Different terminologies to encode the most common signs, symptoms and problematic foods associated with IgE-mediated food allergies were identified. Their suitability for classifying and coding information about the signs and symptoms of food allergic reactions, causative foods and reaction severity of was assessed. The assessment included existing conceptual coverage and data descriptions, classification schemes and additional relevant information.

Results: All of the terminologies reviewed included classification schemes, allowing broader concepts to be related to those that are more specialised. Additional information was often present such as equivalence. Of the clinical coding systems assessed, the Systemized Nomenclature of Medical Clinical Terms (SNOMED-CT) provided the most complete coverage with options to code symptom severity. Only food coding systems, such as FoodEx2, provided comprehensive conceptual coverage of the food terms.

Conclusions: Utilising SNOMED-CT and FoodEx2 standards together will support the harmonisation of data regarding food allergy from diverse sources, providing a transparent and effective way to collate relevant data required for effective food allergen management in the future.

Introduction

IgE-mediated food allergies are estimated to affect around 1% of infants ^{1,2} and up to 4% of adults ³ in Europe. The prevalence varies between countries with rates being higher in Australia ^{4,5} and lower in countries such as India and parts of China ⁶. Approved therapies are currently only available in the USA and are solely used to treat peanut allergy in children and adolescents. The therapy can protect allergic individuals from accidental reactions ⁷ but are not accessible to all and not available for all the different foods that can precipitate an allergic reaction. Consequently, avoidance of the causative food is generally the only recourse and in order to help individuals avoid their problem food, labelling of a group of priority allergenic foods has been made mandatory in most parts of the world. However, traces of allergenic ingredients can find their way, unintentionally, into food products and can cause adverse reactions ⁸. In order to warn allergic consumers of the potential hazard posed by the presence of unintended allergens, food manufacturers place precautionary allergen labels (PAL) on such foods. PAL should only be used in conjunction with a risk assessment ⁹ but such approaches require reference doses of allergens that are accepted as being generally safe for the majority of food allergic consumers and are usually derived from oral food challenge data ¹⁰. Such risk based approaches to food allergen management also provide a transparent and consistent system to prevent the overuse of 'may contain' labelling on foods where

allergens are less likely to pose a risk and avoid mistrust from allergic consumers which can lead to risky choices being made of foods that could result in a reaction.

The ThRAI project aims to support the application of risk-based approaches to food-allergen management ¹¹. This involves the collation, harmonization and integration of data from individuals with IgE-mediated allergies undergoing a diagnostic procedure called an oral food challenge ¹². These can be used to identify doses of allergens below which food allergic subjects are unlikely to react, or react with only mild symptoms ¹³. Through the ThRAI project a publicly available database of oral food challenges is being developed, which can be used for dose distribution modelling that underpins identification of doses that have an acceptable level of risk of eliciting a reaction in the allergic population ¹⁰. Data from oral food challenges are often collected and reported in a heterogeneous manner, with studies conducted in different ways (double blind placebo controlled, single blind [when only the patient is blinded], interspersed, open challenge), using different dosing protocols and food matrices to deliver the allergenic food. In addition, different investigators can use multiple terminologies to describe the same clinical symptom (i.e. a change in function, sensation or appearance which indicates disease which is reported by the patient) or sign (i.e. an objective observation or evidence of disease) even within the same study centre. This ambiguity can lead to different interpretations when harmonizing existing datasets. To facilitate the integration of existing data and support repeatability of future studies it is important to have a consensus approach to encoding and reporting food allergy information. This paper aims to identify and encode common variables relating to food allergy to promote the repeatability of data analysis and facilitate the interoperability of data being collated in the ThRAI project relating to food allergies.

Materials And Methods

Relevant concepts that are used to understand, describe and diagnose food allergies were first identified and defined (Supplementary material Table 1). Clinical record forms used for recording oral food challenges from the EuroPrevall (The prevalence, cost and basis of food allergy across Europe) and iFAAM (Integrated approaches to food allergen and allergy management) studies were then used to identify common signs and symptoms experienced during an IgE-mediated reaction, which were also defined (Supplementary material Table 2) ^{9,14-17}. In addition, key food terms were compiled based on the foods that initiate an allergic reaction (including species of origin, common derivative ingredients and manufactured foods) based on the 14 major allergens which must be stated on food labelling as described by Annex II of the EU Food Information for Consumers Regulation 1169/2011 ^{18,19} (Supplementary material Table 3).

In the next step the utility of different terminologies for classifying and encoding this information was assessed. Clinically relevant and validated terminologies were identified using The National Library of Medicine. These were the Systemized Nomenclature of Medical Clinical Terms (SNOMED-CT), Medical Subject Headings (MeSH) and the Medication Dictionary for Regulatory Activities (MedDRA) systems (Table 1). LOINC (Logical Observation Identifiers Names and Codes) was also considered for classifying

and encoding the symptom concepts. However, this system is used to represent the type or “question” for a clinical test or measurement. This is not in scope for this study, where we focus on coding observation results or “responses”, and LOINC was removed from further consideration. In addition, the European Food Safety Authorities FoodEx2, LanguaL alimentaria (LanguaL) and a vocabulary developed by the Food and Agriculture Organisation (AGROVOC) were compared for their classification of food products (Table 1).

Table 1
Terminologies assessed for their utility in encoding food allergy data.

Terminology	Description
SNOMED-CT (Systemized nomenclature of medicine- clinical terms) ²⁰	SNOMED-CT is the most comprehensive and international clinical terminology system.
MeSH (Medical subject headings) ²¹	MeSH is used for indexing, cataloguing and searching biomedical and health-related information.
MedDRA (Medical Dictionary for Regulatory Activities) ²²	MedDRA provides a clinically validated and internationally recognised medical terminology.
LOINC (Logical Observation Identifiers Names and Codes) ²³	LOINC is an international standard, which facilitates the storage, exchange and harmonization of data.
FoodEx2 ²⁴	FoodEx2 is a standardised food classification and description system. Data is compiled from EU organizations, industries and academic research.
LanguaL (Langua aLimentaria" or "language of food") ²⁵	LanguaL provides a standardised technique for describing foods based on the combination of characteristics of that food ²⁶ .
AGROVOC (Agriculture and Vocabulary) FAO (Food and Agriculture Organization of the United Nations) ²⁷	AGROVOC is a controlled vocabulary including both food and nutrition that can translate concepts into 37 languages.

The ThRAI project aimed to evaluate the classification of symptoms, signs and food relating to an IgE-mediated food allergy from a risk management and public health perspective. Given this objective each terminology was assessed in four areas:

1. *Conceptual coverage*: Every symptom and food term identified was searched in each of the terminologies to quantitatively compare which system provided the greatest existing conceptual coverage.
2. *Concept descriptions*: A clear and validated definition was identified for each symptom or sign and food term. This was used as a ‘benchmark’ definition to be assessed and compared with the descriptions provided in each of the terminologies.
3. *Classification*: The information provided by each terminology to support classification of specific symptoms and signs together with food were compared. The inclusion and clinical appropriateness

of information that expresses the relationships between general and specialised concepts (classification schemes) was considered to be a key differentiator.

4. *Additional information:* Each terminology was assessed for support for synonyms specified in the ThRAI protocol. This was completed for symptoms and signs, but not for food where equivalence is more complex to determine. Any additional information was also assessed for use in the ThRAI study.

Results

Conceptual coverage

For the symptom and sign concepts (Supplementary material Table 1) both SNOMED-CT and MedDRA provided complete coverage, whilst MeSH covered 88%. FoodEx2, LanguaL and AGROVOC did not cover any of the symptom concepts but did provide superior coverage for food concepts (Table 2). AGROVOC covered 85% of the food concepts, LanguaL had complete coverage and FoodEx2 covered all of the food concepts except for sulphur dioxide. In addition, LanguaL uses the EFSA FoodEx2 coding and classification system for products in the European Union, which validates the legitimacy and effectiveness of FoodEx2.

Table 2

Comparison of five terminologies for the classification and coding of food allergy information.

Terminology	Conceptual coverage		Classification	Concept description	Additional information
	Symptom (n = 25)	Food (n = 47)			
SNOMED-CT	25	42	Concepts are arranged into a hierarchical structure.	Detailed and unambiguous definition.	Preferred term and synonyms.
MeSH	22	32	Concepts are arranged into a logical and detailed hierarchical structure.	Detailed and unambiguous definition.	Provides related concepts (where available) for each term.
MedDRA	25	6	Terms are arranged into a 5 tier branching system expanding from very specific to more general concepts.	No definition provided.	Includes a preferred name and related synonyms.
FoodEX2	0	46	Concepts are categorised into one of 21 groups and facets can encode additional detail including ingredients or processing techniques.	Clear description along with common and scientific name	N/A
LanguaL	0	47	Concepts are classified systematically based on 14 key terms including product type, food source and cooking method.	Clear description along with common and scientific name	N/A
AGROVOC	0	40	Terms are arranged in a hierarchical and non-hierarchical system to classify a range of concepts and to indicate related terms.	No description is provided, but the broader concept is included to provide context.	A list of related concepts are included for some terms.

Description of concepts

The term definitions provided by each of the coding systems were reviewed and compared with the definitions found in the literature (Supplementary material Tables 2 and 3). MedDRA does not provide formal definitions for the symptom and sign concepts and so fails to provide any additional clarity or description for each term. In contrast, SNOMED-CT and MeSH provide clear and unambiguous descriptions for each concept. This is illustrated in Table 3 using “urticarial rash” and “hazelnut” as an

example sign and food respectively. Thus, both MeSH and SNOMED-CT include a detailed definition to describe the sign “Urticarial rash” which is consistent with the key definitions (Supplementary material Table 2). Similarly, FoodEx2, LanguaL and AGROVOC all provide a detailed description for each food term as well as detailing both the common and the Latin name for each food to reduce ambiguity (Supplementary material Table 3).

Table 3. Comparison of the definition of exemplar symptom (urticarial rash) and food (hazelnuts) terms according to the different terminologies.

Terminology	Definition
Urticarial rash	
SNOMED-CT	“A raised, erythematous papule or cutaneous plaque usually representing short-lived dermal oedema.”
MeSH	“A vascular reaction of the skin characterised by erythema and wheal formation due to localized increase of vascular permeability. The causative mechanism may be allergy, infection or stress.”
MedDRA	N/A
ThRAIL approach	“A condition characterized by the development of wheals (hives), angioedema or both”
Hazelnut	
SNOMED-CT	“Tree nut (substance)”
MeSH	“A plant genus of the family BETULACEAE known for the edible nuts”
FoodEX2	“Tree nuts from the plant classified under the species <i>Corylus avellana</i> L., commonly known as Hazelnuts or Cobnuts or Common hazelnut. The part consumed/analysed is not specified. When relevant, information on the part consumed/analysed has to be reported with additional facet descriptors. In case of data collections related to legislations, the default part consumed/analysed is the one defined in the applicable legislation.”
LanguaL	“The group includes kernels of the seeds of all species similar to Hazelnuts or similar nuts sharing the same pesticide to the maximum residue level (MRL) as Hazelnuts.”
AGROVOC	“The fruit of small trees or shrubs of the Corylaceae family. The round-oval nuts are surrounded by a leafy involucre, which comes out easily when the fruit is mature. Remains the nut, with a pericarp, the shell, more or less woody, and depending on varieties. Inside is the seed, covered by a very thin tegument.”

Where appropriate, MedDRA aggregates and highlights similar terms related to the specific concept and so allows comparable terms to be easily accessed. For example, ‘urticaria rash’ is the preferred name that classifies 33 related concepts including ‘urticaria localized’ and ‘generalised urticarial rash’, which provide

varying levels of detail and alternative terms that include a level of clinical interpretation. This is useful but the large number of similar terms may reduce the specificity and level of detail initially identified by a term. In contrast, concepts in SNOMED-CT are associated with a unique Fully Specified Name (FSN); this is the ideal term that a clinician would use in a particular language, dialect or context. SNOMED-CT also identifies relationships to other similar and related concepts and considers the preferred name for 'urticarial rash' to be 'weal' with 'wheal', 'welt', 'hives' and 'nettle rash' being noted as alternative terms. Since concepts are coded in MeSH for the purpose of indexing publication records, each MeSH term can comprise several synonyms; for example, 'urticaria' also includes 'urticarias' and 'hives'. All concepts that come under one record are considered equivalent and this is useful when trying to maximise the number of relevant articles identified in a search but not necessarily when identifying synonyms in the context of compiling data on food allergy in the ThRAI database.

LanguaL and FoodEx2 do not provide related concepts since this is not appropriate in the context of a food classification system as there would be no suitable synonyms. In certain cases AGROVOC provides equivalent terms and where appropriate states the broader and/or narrower relevant concepts as well as identifying what the product is produced (i.e. the plant/ animal species of origin).

Mechanistic basis to classify an adverse reaction

It was also important to compare the pathway that each of the terminologies used to classify an adverse reaction. Food can induce a range of adverse reactions but, although the symptoms and signs may be similar, the mechanistic basis of allergies, intolerances and sensitivities are completely different. An adverse reaction to food ²⁸ encompasses both immune- and non-immune mediated adverse reactions (Fig. 1). Sub-types of immune-mediated adverse reactions include those involving the development of food-specific IgE antibody responses. This type of food allergy results in symptoms and signs that appear immediately (in less than 2 hours, usually within 30 minutes) sometimes even after ingesting a small dose of the allergen, and can involve multiple organs including the skin, respiratory, digestive and cardiovascular systems. A second type of well-defined non-IgE-mediated adverse reaction to food is the T-cell mediated syndrome triggered by ingestion of gluten, known as coeliac disease (CD)²⁹. This life-long disease involves sensitivity to gluten and individuals with CD often present with gastrointestinal signs and symptoms, including diarrhoea, together with weight loss due to the malabsorption of nutrients. In contrast, food intolerance conditions are not immune mediated but nevertheless can be reproducibly induced following ingestion of specific foods. One example is lactose intolerance where individuals lack the lactase enzyme, which is involved with the digestion of lactose. Symptoms appear shortly after drinking milk or consuming dairy and are commonly reported as stomach pain, bloating and diarrhoea. Lactose intolerance is different to an IgE mediated milk allergy, which is an IgE mediated reaction where symptoms appear within 2 h of consuming milk-containing foods.

The classification of an adverse reaction used in the ThRAI project (Fig. 1) was used to benchmark how the different terminologies classify an IgE mediated food allergy (Fig. 2). SNOMED-CT considers the trigger of an allergic reaction as the causative food and then links this back to the allergic

hypersensitivity. It provides an overview of the process of the reaction but also filters into specific details about the adverse response, including branching to the causative agent and a qualifier for the severity of the reaction. This is consistent with terminology from the World Allergy Organisation (WAO) and the European Academy of Allergy and Clinical Immunology (EAACI)³⁰. MedDRA and MeSH classify food allergy from a disease perspective and then acknowledge the response as a consequence from ingestion of the problematic food. MeSH also includes a logical and clear flow using the descriptor “food hypersensitivity” as a type of “immediate hypersensitivity” which is used as a synonym of IgE-mediated food allergy. This is linked to the causal food and the eliciting symptoms. In contrast, MedDRA utilizes a tree flow diagram to show how a food allergy is classified but this was a simpler and less detailed pathway compared to MeSH and SNOMED-CT. Since the ThRAI project is considering the reaction from a food and public health perspective, the SNOMED-CT approach to classify a food allergy was considered the most appropriate.

Figure 3 illustrates how MeSH, MedDRA and SNOMED-CT classify a non-IgE immune mediated adverse reaction to food and coeliac disease. MedDRA and MeSH both provide multiple ways to classify the pathway of coeliac disease. These terminologies consider this disease as a nutritional or gastrointestinal disorder that leads to malabsorption, which clearly demonstrates the reaction is triggered by food. MedDRA has a third pathway to classify coeliac disease from a disease perspective as an autoimmune disorder; this is consistent with the ThRAI approach, which classifies coeliac disease as an immune-mediated reaction. Again, SNOMED-CT has a different way of approaching the classification of coeliac disease in comparison to MedDRA and MeSH but still considers it a malabsorption syndrome caused by the ingestion of gluten. Figure 3 demonstrates that coeliac disease is an adverse response with a clear food trigger, yet the pathway to classify this reaction is very different to the classification of a food allergy and so these reactions should be considered separately.

Lastly the pathways to describe a non-immune mediated adverse reaction by specifically looking at the classification of lactose intolerance are shown in Fig. 4. MedDRA, MeSH and SNOMED-CT all describe lactose intolerance as a metabolic or gastrointestinal disorder which disrupts the absorption of carbohydrates. The ThRAI approach classifies lactose intolerance as a non-immune mediated reaction due to a disorder of an enzymatic process. The lack of lactase enzyme in lactose intolerance patients causes the malabsorption of the carbohydrate lactose and so demonstrates the similarities between these classifications.

Classification schemes for symptoms and signs

MedDRA has a logical five-tier structure expanding from very specific to more general concepts. Lowest Level Terms (LLTs) represent the most specialised concepts including symptoms, medical procedures and personal characteristics. LLTs can be considered the preferred term (PT) or a synonym of the preferred term (PT)³¹. Similar PTs are aggregated into High Level Terms (HLT) based upon anatomy, pathology, physiology or aetiology. These HLTs are further categorised into High Level Group Terms (HLGT) that are then split into one of 26 System Organ Classes (SOC's) providing the most general

classification. This is a logical and methodical organisation system but concepts are confined to these five levels and further clarification or granularity cannot be expressed beyond the LLT to indicate the severity or manifestation of a symptom. Concepts in SNOMED-CT can vary in their specificity; more general concepts are aggregated together which filter down to more specific terms. Relationships are used to portray a confirmed association between multiple concepts. The branching structure in SNOMED-CT is useful to be able to code and represent clinical data at a level of detail that is appropriate to a range of different uses. In contrast MeSH is a cataloguing system with a slightly different framework to the other terminologies as terms do not identify clinical phenomena but instead represent a category. MeSH is used to categorise and retrieve records and is organised in a hierarchical structure with 16 primary categories that splits into subcategories providing more detailed terms. A letter corresponding to a category and a number representing the hierarchical level provides an identifier for each term.

Classification schemes for food

There are many levels of granularity that need to be considered when describing and encoding a food product including the origin of the food, the food matrix and processing techniques. For example, when considering a food allergy, we need to identify differences in the frequency or severity of a reaction,¹ which may be affected by cooking technique (e.g. dry roasted compared to raw peanuts) or the food matrix/ vehicle used (e.g. peanut butter, whole peanut, baked goods containing peanuts). It is useful to have a system with the ability to encode multiple levels of detail depending on the amount of data and information that is available. This is important given that the literature shows that some food allergens are sensitive to food-processing techniques and a high fat content may increase the allergenicity of the protein ^{32,33}. The structure of these different classification systems represents their primary use whether that is to understand nutritional value, physical characteristics or the type of food product.

FoodEx2 is arranged into 21 clearly defined food groups, such that every food aligns to exactly one group. The system is made up of base terms and facets; the base term is defined by a unique five-character alphanumeric code and represents the specific foods within the hierarchy. Facets provide additional detail to the base term, such as the origin of the product; its ingredients and the process involved in its preparation. This additional information can be combined with the base term to provide a more detailed and complete description of the food product. This demonstrates the range of granularity and amount of information that has been considered in this classification system making it useful for encoding allergenic foods and common matrices or derivatives used in oral food challenges.

LanguaL systematically classifies food in a systematic way according to 14 key concepts including product type, cooking method and packing medium. These concepts are used to encode the product and enable almost any food product to be classified to the level of detail that is required. A unique code is provided for each food concept, which can then be translated into multiple languages.

AGROVOC relates concepts in a hierarchical and non-hierarchical structure. A branching logic is used whereby the term becomes more specific and precise. For example, nuts provides a broader way of

describing hazelnuts. There is also a non-hierarchical relationship and this expresses related concepts where appropriate.

Figure 6 demonstrates how each of the standards classifies hazelnuts as an example of a food concept. The figure shows that LanguaL utilises the logic from FoodEx2 and that each of the terminologies distinguish between plant and animal products before identifying the relevant concept from a list of key categories and then filtering into the specific species.

Encoding symptom severity

When considering the dose at which an allergen induces an objective reaction it is useful to encode the severity of that reaction. Thus, when identifying the dose that elicits a reaction in p% of the allergenic population (eliciting dose, ED_p), it is useful to understand the proportion of individuals that presented with either a mild, moderate or severe reaction at this dose. Some studies, such as iFAAM, have collected detailed information on severity which can also be used to identify criteria for stopping a challenge [16]. This information can be used to identify eliciting doses that present a tolerable level of risk of reaction in allergic individuals, as observed in the Peanut Allergen Threshold study ¹³. The classification of reaction severity is inherently subjective, arising from clinical interpretation of signs and symptoms and many different approaches have been developed and compared ^{34,35}. SNOMED-CT codes do not always provide sufficient detail to describe the severity as well as symptom presence but this can be addressed by using an additional qualifier code (mild, moderate or severe) to fully express the specific, observed sign or symptom. This is illustrated for the iFAAM oral food challenge record where symptoms have been classified and coded with both the symptom code and the terms, mild, moderate or severe using the approach of Sampson ³⁶ (Table 4). This shows, for example, if a patient experienced one episode of diarrhoea this would be coded as the SNOMED-CT code for diarrhoea and the code of mild severity and paired in a structure to indicate one episode of diarrhoea: 62315008–255604002.

Table 4
Encoding severity alongside symptoms and signs using SNOMED-CT.

Symptom	SNOMED-CT code	Description for relevant severity grade (SNOMED-CT code)		
		Mild (255604002)	Moderate (6736007)	Severe (24484000)
Pruritus	418363000	<ul style="list-style-type: none"> Occasional scratching Continuous scratching for > 2 min at a time 	<ul style="list-style-type: none"> Hard continuous scratching leading to excoriations Scratching of palms, soles, genitals, scalp 	
Erythema	247441003	<ul style="list-style-type: none"> Few areas of faint erythema Areas of erythema (< = 50%) 	Generalised marked erythema (> 50%)	
Urticarial rash (wheal)	247472004	Up to 10 new hives	Generalised involvement (> 10 new hives)	
Angioedema	41291007	Mild lip oedema	<ul style="list-style-type: none"> Significant lip oedema Whole face oedema 	
Rhinitis	70076002		<ul style="list-style-type: none"> Rare bursts, occasional sniffing < 10 bursts, frequent sniffing or intermediate rubbing of nose; Long bursts, persistent rhinorrhoea or continuous rubbing 	
Ocular		Intermittent rubbing of eyes	Continuous rubbing, periocular swelling, reddening	

Signs and Symptoms recorded in the iFAAM challenge record [16] were classified and coded as to their severity using a combination of the approach of Sampson ³⁶ as being either mild (Sampson grade 1), moderate (Sampson grades 2 and 3) or severe (Sampson grades 4 and 5) and then encoded using SNOMED-CT.

Symptom	SNOMED-CT code	Description for relevant severity grade (SNOMED-CT code)		
		Mild (255604002)	Moderate (6736007)	Severe (24484000)
Wheezing	56018004; 272040008			<ul style="list-style-type: none"> • Expiratory wheezing to auscultation • Inspiratory and expiratory wheezing to auscultation • Use of accessory muscles or audible wheezing
Gastrointestinal pain Nausea	21522001 422587007		<ul style="list-style-type: none"> • Complaints of nausea or abdominal pain • Frequent complaints of nausea or pain with abnormal activity • Notably distressed due to GI symptoms with decreased activity 	
Emesis	422400008		<ul style="list-style-type: none"> • 1 episode of emesis • > 1 episode of emesis 	
Diarrhoea	62315008			<ul style="list-style-type: none"> • 1 episode of diarrhoea • > 1 episode of diarrhoea
Signs and Symptoms recorded in the iFAAM challenge record [16] were classified and coded as to their severity using a combination of the approach of Sampson ³⁶ as being either mild (Sampson grade 1), moderate (Sampson grades 2 and 3) or severe (Sampson grades 4 and 5) and then encoded using SNOMED-CT.				

Discussion

The results demonstrate the complexity required to encode food allergy information and each terminology has a different way of classifying an allergic reaction. Furthermore, each terminology varies in the system used to classify and organise concepts, which depend on the primary purpose of the classification system. The analysis showed that SNOMED-CT encoded all of the key terms and provided the conceptual coverage required to maximise the representation and classification of food allergy data for public health purposes aimed at food allergen management. In addition, SNOMED-CT provides a classification scheme that relates specific and more general concepts and that considers food as the trigger of an allergic reaction which is clinically relevant. Therefore, it was chosen as the most appropriate

terminology to code data relating to food allergies collated within the ThRAI project. Additionally, SNOMED-CT also provides severity code, allowing for more precise, qualified records of reactions.

Although FoodEx2, SNOMED-CT, MeSH, LanguaL and AGROVOC all provide terms for foods, FoodEx2 provided the greatest conceptual coverage including species of origin and food processing methods. Consequently, FoodEx2 was selected to code the foods that initiate an allergic reactions, enabling analysis to be undertaken to identify how factors, such as food processing, may influence eliciting dose threshold and reaction severity.

Conclusion

The ThRAI study has determined that SNOMED-CT and FoodEx2 are presently the most competent terminologies for the representation of clinical knowledge relating to food allergy, the former delivering encodings for observed symptoms or signs and the latter representing eliciting foods (Fig. 6). This combined set of terminologies provides not only complete conceptual coverage, but a depth of granularity providing a structured flexibility to encoding a given required level of detail, via symptom severity in SNOMED-CT, and facets for food matrices and processing techniques in FoodEx2.

It is expected that by defining a coding system for the ThRAI oral food challenge database will increase accessibility through standardisation. This will in turn extend the usefulness of these data into the future. Furthermore, the approach taken could also be expanded in future to encompass data from other clinical studies and registries of food anaphylaxis such as the network of severe allergic reactions (Network for Online-Registration of Anaphylaxis, NORA) ³⁷. The proposed pairing of codes to qualify severity of reaction is supported in HL7's FHIR (Fast Healthcare Interoperability Resource), a standard for exchanging healthcare information electronically. This system represents data as a set of related hierarchical elements and values into resources. For the purpose of capturing data relating to allergic reactions, it is possible to record a symptom element, a severity element and a food element within the same resource meaning that there is a standard definition of the relationships between the codes. Using the proposed approach to coding within this standard could further increase the future utility of these data making for a promising approach for recording and storing reusable datasets. The coding approach is also applicable to representing an individual's food allergy information in a standardised way in their electronic health records by including other metadata, such as diagnostic certainty (e.g. whether the data was a self-reported, negative, possible, probable or confirmed allergy) which could ensure correct interpretation across medical systems.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

The authors declare no competing interests in relation to the published work.

Funding

This project has received financial support from the European Food Safety Authority (EFSA), Grant GP/EFSA/AFSCO/2017/03. The present article, however, is under the sole responsibility of the authors. The positions and opinions presented in this article are those of the authors alone and do not necessarily represent the views/any official position or scientific works of EFSA. EFSA guidance documents and other scientific outputs of EFSA, can be found in the EFSA website: <http://www.efsa.europa.eu>. The ThRAI project is also co-funded by the UK Food Standards Agency FS101209. This work was partly funded by the European Union through the iFAAM project: Integrated Approaches to Food Allergen and Allergy Risk Management (Grant Agreement N° 312147) and the Medical Research Council Health eResearch Centre, Farr Institute (MR/K006665/1).

Authors' contributions

PC and ENCM were responsible for the original concept and design of the work and have overall supervision of its execution. CF, BJ, BG and S L-T were responsible for its execution. MM, AK, T-ML, AS, SD, NdJ, BB-W, MF-R and KB were involved in providing clinical oversight and revision of the work. The manuscript was drafted by CF, BJ, ENCM and PC and all authors were involved in its review and revision and final approval.

References

1. Schoemaker AA, Sprickelman AB, Grimshaw KE, et al. Incidence and natural history of challenge-proven cow's milk allergy in European children–EuroPrevall birth cohort. *Allergy*. 2015;70(8):963-972.
2. Xepapadaki P, Fiocchi A, Grabenhenrich L, et al. Incidence and natural history of hen's egg allergy in the first 2 years of life-the EuroPrevall birth cohort study. *Allergy*. 2016;71(3):350-357.
3. Lyons SA, Burney PGJ, Ballmer-Weber BK, et al. Food Allergy in Adults: Substantial Variation in Prevalence and Causative Foods Across Europe. *The journal of allergy and clinical immunology In practice*. 2019;7(6):1920-1928.e1911.
4. Osborne NJ, Koplin JJ, Martin PE, et al. Prevalence of challenge-proven IgE-mediated food allergy using population-based sampling and predetermined challenge criteria in infants. *J Allergy Clin Immunol*. 2011;127(3):668-676 e661-662.

5. McWilliam V, Peters R, Tang MLK, et al. Patterns of tree nut sensitization and allergy in the first 6 years of life in a population-based cohort. *The Journal of allergy and clinical immunology*. 2019;143(2):644-650.e645.
6. Li J, Ogorodova LM, Mahesh PA, et al. Comparative Study of Food Allergies in Children from China, India, and Russia: The EuroPrevall-INCO Surveys. *The journal of allergy and clinical immunology In practice*. 2019.
7. Trendelenburg V, Blumchen K, Bellach J, et al. Peanut oral immunotherapy protects patients from accidental allergic reactions to peanut. *The journal of allergy and clinical immunology In practice*. 2020.
8. Blom WM, Michelsen-Huisman AD, van Os-Medendorp H, et al. Accidental food allergy reactions: Products and undeclared ingredients. *J Allergy Clin Immunol*. 2018;142(3):865-875.
9. DunnGalvin A, Roberts G, Schnadt S, et al. Evidence-based approaches to the application of precautionary allergen labelling: Report from two iFAAM workshops. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*. 2019;49(9):1191-1200.
10. Allen KJ, Remington BC, Baumert JL, et al. Allergen reference doses for precautionary labeling (VITAL 2.0): clinical implications. *J Allergy Clin Immunol*. 2014;133(1):156-164.
11. Mills ENC, Adel-Patient K, Bernard H, et al. Detection and Quantification of Allergens in Foods and Minimum Eliciting Doses in Food-Allergic Individuals (ThRAII). *J AOAC Int*. 2019;102(5):1346-1353.
12. Sampson HA, Gerth van Wijk R, Bindslev-Jensen C, et al. Standardizing double-blind, placebo-controlled oral food challenges: American Academy of Allergy, Asthma & Immunology-European Academy of Allergy and Clinical Immunology PRACTALL consensus report. *The Journal of allergy and clinical immunology*. 2012;130(6):1260-1274.
13. Hourihane JO, Allen KJ, Shreffler WG, et al. Peanut Allergen Threshold Study (PATS): Novel single-dose oral food challenge study to validate eliciting doses in children with peanut allergy. *J Allergy Clin Immunol*. 2017;139(5):1583-1590.
14. Kummeling I, Mills EN, Clausen M, et al. The EuroPrevall surveys on the prevalence of food allergies in children and adults: background and study methodology. *Allergy*. 2009;64(10):1493-1497.
15. Fernandez-Rivas M, Barreales L, Mackie AR, et al. The EuroPrevall outpatient clinic study on food allergy: background and methodology. *Allergy*. 2015;70(5):576-584.
16. Grabenhenrich LB, Reich A, Bellach J, et al. A new framework for the documentation and interpretation of oral food challenges in population-based and clinical research. *Allergy*. 2016.
17. Grabenhenrich LB, Reich A, Bellach J, et al. A new framework for the documentation and interpretation of oral food challenges in population-based and clinical research. *Allergy*. 2017;72(3):453-461.
18. Commission E. Food information to consumers- Legislation 2016;
https://ec.europa.eu/food/safety/labelling_nutrition/labelling_legislation_en. Accessed 17/08/2020.
19. [Annex II. FESUM internal regulations]. *Chir Main*. 2003;22(5):267-269.

20. Digital N. SNOMED-CT Browser. 2017; v1.36.4:<https://termbrowser.nhs.uk/?perspective=full&conceptId1=404684003&edition=uk-edition&release=v20200805&server=https://termbrowser.nhs.uk/sct-browser-api/snomed&langRefset=999001261000000100,999000691000001104>. Accessed 18/08/2020.
21. Medicine USNLo. Medical Subject Headings 2020. 2020; <https://meshb.nlm.nih.gov/search>. Accessed 18/08/2020.
22. Activities MDfR. Medical Dictionary for Regulatory Activities. 2019.
23. LOINC. LOINC 2.68:<https://loinc.org/>.
24. Authority EFS. The food classification and description system FoodEx2 (revision 2). 2015; <https://www.efsa.europa.eu/en/supporting/pub/en-804>. Accessed 18/08/2020.
25. TM L. The LanguaL 2017 Thesaurus Systematic Display. 2017; https://www.languaL.org/languaL_Thesaurus.asp. Accessed 18/08/2020.
26. Ireland JD, Møller A. LanguaL food description: a learning process. *European journal of clinical nutrition*. 2010;64 Suppl 3:S44-48.
27. Nations FaAOotU. AGROVOC Multilingual Thesaurus. 2020; <http://aims.fao.org/standards/agrovoc/functionalities/search>. Accessed 18/08/2020.
28. Johansson SG, Hourihane JO, Bousquet J, et al. A revised nomenclature for allergy. An EAACI position statement from the EAACI nomenclature task force. *Allergy*. 2001;56(9):813-824.
29. Pes GM, Bibbò S, Dore MP. Coeliac disease: beyond genetic susceptibility and gluten. A narrative review. *Annals of medicine*. 2019;51(1):1-16.
30. Cartagena FP, Schaeffer M, Rifai D, Doroshenko V, Goldberg HS. Leveraging the NLM map from SNOMED CT to ICD-10-CM to facilitate adoption of ICD-10-CM. *J Am Med Inform Assoc*. 2015;22(3):659-670.
31. Goss FR, Zhou L, Plasek JM, et al. Evaluating standard terminologies for encoding allergy information. *J Am Med Inform Assoc*. 2013;20(5):969-979.
32. Grimshaw KE, King RM, Nordlee JA, Hefle SL, Warner JO, Hourihane JO. Presentation of allergen in different food preparations affects the nature of the allergic reaction—a case series. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*. 2003;33(11):1581-1585.
33. Mackie A, Knulst A, Le TM, et al. High fat food increases gastric residence and thus thresholds for objective symptoms in allergic patients. *Molecular nutrition & food research*. 2012;56(11):1708-1714.
34. Eller E, Muraro A, Dahl R, Mortz CG, Bindslev-Jensen C. Assessing severity of anaphylaxis: a data-driven comparison of 23 instruments. *Clinical and translational allergy*. 2018;8:29.
35. Muraro A, Fernandez-Rivas M, Beyer K, et al. The urgent need for a harmonized severity scoring system for acute allergic reactions. *Allergy*. 2018;73(9):1792-1800.
36. Sampson HA. Anaphylaxis and emergency treatment. *Pediatrics*. 2003;111(6 Pt 3):1601-1608.

37. Worm M, Moneret-Vautrin A, Scherer K, et al. First European data from the network of severe allergic reactions (NORA). *Allergy*. 2014;69(10):1397-1404.

Figures

Figure 1

Classification of a non-toxic adverse reactions to food.

Figure 2

Comparison of the pathways to describe the onset of IgE-mediated food allergy according to MedDRA, MeSH, SNOMED-CT and the ThRAI approach.

Figure 3

Comparison of the pathways to describe Coeliac disease, a non-IgE immune mediated adverse reaction to food, according to MedDRA, MeSH, SNOMED-CT and the ThRAI approach.

Figure 4

Comparison of the pathways to describe lactose intolerance, a non-immune mediated adverse reaction to food, according to MedDRA, MeSH, SNOMED-CT and the ThRAIL approach.

ONTOLOGY

Figure 5

Branching logic to classify the sign “urticarial rash” according to MedDRA, MeSH and SNOMED-CT. *SNOMED-CT preferred tem is Wheal (disorder) with urticarial rash as a synonym.

ONTOLOGY

Figure 6

Branching logic to classify hazelnuts according to FoodEx2, LanguaL and AGROVOC.

Figure 7

The ThRAII pathway for coding food allergy information using SNOMED-CT and FoodEX2

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FrenchetalFoodAllergyCodingSupplInformationver8.docx](#)