Title: Nutrition Informatics (Nutri-informatics): A call for improved and integrative standards for nutrition research

Authors: Lauren Chan[1], Nicole Vasilevsky[2], Anne Thessen[3], Julie McMurry[1], and Melissa Haendel[2,3]

1: College of Public Health and Human Sciences, Oregon State University
2: Oregon Health and Science University
3: Environmental and Molecular Toxicology Department, Oregon State University

Abstract:

Nutri-informatics aims to computationally integrate and analyze nutrition study datasets in order to disentangle the interactions between an organism and its nutritional environment. Fueled by an interest in how food, nutrients, and nutrition sociology impact health, and a recent push towards “big data”, nutri-informatics is essential to incorporating nutrition into computational biomedical sciences.

Nutrition is one of the most integral components to human life, and it impacts individuals far beyond just nutrient provisions. For example, nutrition plays a role in cultural practices, interpersonal relationships, and body image. Despite this, integrated computational investigations have been limited due to challenges within nutri-informatics and nutrition data.

Nutri-informatics suffers from a lack of standardization with a wide array of groups working on similar projects with no community-wide development principles to ensure interoperability and cohesion between nutri-informatics and other biomedical resources. While a large number of resources for nutri-informatics are available, much of nutrition is underrepresented. This may be due to how expansive and heterogeneous nutrition is as a field, increasing the difficulty of data modeling. Approaches to formalize nutrition research language and connect standardized terminologies across biomedical fields have been initiated through the use of biomedical ontologies and computational nutrition data resources. While a variety of nutrition-related ontologies such as Food Ontology and the Ontology for Nutritional Studies have been initiated, they are still in development and require further attention from nutrition researchers and biomedical ontologists.

Should nutrition data continue to be produced with no standardization of language, documentation specifications, or requirements for data reuse, nutri-informatics investigations will continue to struggle with incompatible data. In efforts to support nutri-informatics, the community must encourage standards for nutrition data production, reuse, and publication. Academic journals as well as members from nutrition research and biomedical ontology communities should promote standardization of language, data interoperability, and FAIR principles.

Nutrition research encompasses a broad swath of human biology

While nutrition and diet are arguably some of the most vital aspects of a healthy life, the study of nutrition as a science is relatively new. Modern day nutrition research began less than 100 years ago with the first vitamin isolation in 1926[1], but has grown into a vast discipline. From a biological standpoint, nutrition is essential to most all living organisms. Life, functions, and reproduction of humans and other organisms are supported by essential nutrients such as water, macronutrients, vitamins, and minerals, obtained from food and drink. Thus, nutrition research has focused on understanding what nutrients are essential[2–4], what foods contain those nutrients[5,6], what biological functions a nutrient may participate in[7,8], how food processing impacts nutrient content[9–11], and evaluation of ideal nutrient needs for individuals with specific health conditions[12–14].

Evidence-based nutrition research has informed clinical practice, such as using food additives to prevent developmental birth defects (e.g. adding folic acid to grain products to reduce the incidence of neural tube defects) or treating impaired metabolic function with nutritional therapy (e.g. consistent carbohydrate diets for glycemic control). Fine tuning individual and population recommendations has been a consistent focus throughout the development of the field of nutrition, for disease prevention and management, and health optimization[15,16]. However, recent advances in understanding nutrient-nutrient interactions[17,18], food- drug interactions[19], molecular processes, and the impact of the microbiome[20,21], make nutrition far more complex than initially thought.

Beyond biochemical investigations, nutrition is distinct in its translation from research to practice as food is personally and culturally rich. While the first and foremost purpose of food for humans is to fulfill the biological need for energy and nutrients, the nature of food intake has biological and cultural cues. Food choices and preparation, number of meals per day, time of eating, and method of eating[22], religious observation, and personal food beliefs[23–25] are just a few examples of how a culture or custom may guide nutritional intake. Access and management of resources can also impact food selection and consumption, as individuals may have limited access to nutritious and/or preferred food items based on location and transportation needs[26]. Individuals with limited monetary resources are also forced to make decisions between food and other necessities such as housing, which can further impact health and safety[27]. The sociological implications of food greatly impact an individual or population’s nutritional intake and quality of life in ways that are not captured from a purely biochemical point of view. This complex nature of food and nutrition creates a highly variable notion as to what ideal nutrition is, while also showcasing how integral food and nutrition are to human daily life and biological function. Due to the deep complexity of nutrition, discussion of health outcomes involving nutrition are arguably incomplete without the inclusion of sociological information. The broad biological, behavioral, and resource driven scope of nutrition and nutrition research is illustrated in Figure 1, depicting how broad categories of nutrition are all interconnected by subcategories. Due to the interrelated nature of nutrition as a whole, nutrition data and research must also be managed in a unified fashion.

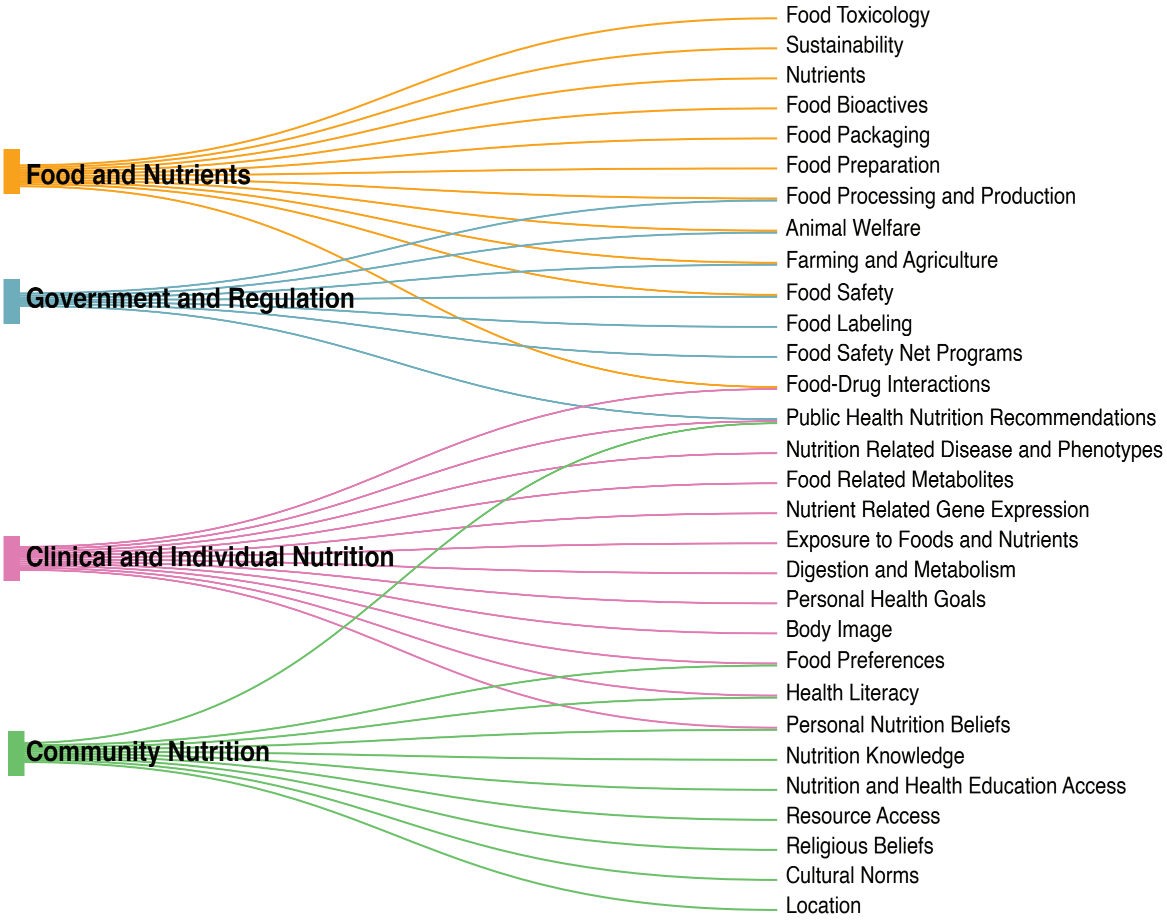


Figure 1. The nutri-informatics landscape. Nutrition is complex and heterogeneous in nature, ranging from larger categories of ‘Food and Nutrients’ to ‘Government and Regulation’, yet within each broad category, many subcategories are shared.

Because nutrition is interdisciplinary and heterogeneous, it is an emergent area for the application of informatics, particularly due to the recent increases in data production through various -omics based nutrition research. **The term nutri-informatics describes approaches to understand the interactions between an organism and its nutritional environment via bioinformatics-based integration of nutrition study data sets**[28]. The desire to utilize informatics approaches to interpret nutrition data can be guided by the successful use of integrated informatics approaches in other biomedical fields, such as genomics, transcriptomics, and metabolomics in combination with more traditional epidemiological and statistical approaches. Currently, nutrition data ranges widely including for example survey data, clinical data, basic science mechanism data, observational data, and -omics data.

Nutri-informatics progress towards improved disease management and precision health

While nutri-informatics may appear to be a new trend, the application of nutri-informatics using advanced statistics has been pursued in nutrition research for some time within large scale investigations of dietary intake via surveys. Surveys such as What We Eat In America (WWEIA), a subset of the National Health and Nutrition Examination Survey (NHANES), are collected annually from Americans in efforts to depict nutritional intake and correlate it with biological samples and clinical measures collected via NHANES[29]. Since the initiation of WWEIA in the 2003-2004 survey period, investigators have capitalized on access to these nutrition data sets for research. Projects using WWEIA and NHANES range widely, and offer insights such as the cost and energy intake associated with dairy replacement in individuals who do not consume dairy products[30], and the prevalence of probable undiagnosed celiac disease and potential reduction in femur bone mineral density[31]. The Nurses' Health Study (NHS), now in its third iteration, similarly collects longitudinal dietary information within their large cohorts via semiquantitative food frequency questionnaires[32]. Since its initiation, NHS dietary information correlated with biological specimen and clinical outcomes in participants have been used to evaluate potential biomarkers for nutrition, such as being the first investigation to measure intake of selenium via toenail samples[32]. Furthermore, NHS has also informed dietary guidelines, such as the recommendation to reduce or eliminate trans fatty acids from the diet to reduce coronary heart disease[32] and highlighting the correlation between eating patterns such as Dietary Approaches to Stop Hypertension and prevention of colorectal cancer in men[33]. Nutrition surveys such as WWEIA and NHS have enabled epidemiological nutrition evaluations with advanced statistics and correlation to clinical and biospecimen data to support improved public health recommendations. While survey-based investigations continue to produce nutrition data to support epidemiological research, approaches to gathering related data have expanded to new data types and a requirement for new methods to incorporate data sources for analysis and inference.

Nutri-informatics initiatives often fall into the category of clinical nutrition, with researchers utilizing informatics to integrate data from electronic health records (EHRs), patient surveys, wearable devices, mobile applications, and other tools to facilitate inference of optimal health recommendations. The Academy of Nutrition and Dietetics, the largest professional group of registered dietitian nutritionists (RDN) in the United States, is striving to participate in the development of standards and processes using nutri-informatics to facilitate optimal nutrition care[34]. This has included support for transitions to EHRs as well as standardization of the electronic Nutrition Care Process and Terminology (eNCPT), a systematic terminology that describes nutrition patient care through Assessment, Diagnosis, Intervention, Monitoring, and Evaluation[34,35]. eNCPT also integrates with the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT) and Logical Observation Identifiers, Names and Codes (LOINC), two commonly used medical terminologies[34]. Implementation of eNCPT in care settings has documented improved efficiency and increased nutrition related diagnoses in hemodialysis patients compared to manual paper based systems, supporting greater effectiveness in patient outcomes[36]. Another approach includes malnutrition identification within a hospital setting[37]. Malnutrition is an extreme risk for hospitalized patients, exacerbating chronic and acute health conditions such as reduced immune function, impaired wound healing and potentially increasing morbidity and mortality rates[37]. Software-based screening tools that standardize malnutrition assessments have improved the consistency and efficiency with which malnutrition is diagnosed, expediting the nutrition care response for patients[37], and also supported the malnutrition knowledge, attitudes, and practices of staff[38].

Nutri-informatics has also been applied in the context of personalized nutrition, i.e. an individual's personal diet and how it translates to health and wellbeing[39]. As both health and disease are highly variable based on genetics, lifestyle, environmental exposures, and many other factors, researchers are focusing on inclusive approaches to develop data-driven predictive methods for anticipating an individual's response to food[39]. An investigation into personalized nutrition for glycemic control by Zeevi et al. utilized machine learning techniques tracking anthropometrics, dietary intake, individual microbiome, and glycemic status to develop a predictive model for postprandial glucose response (PPGR)[13]. These findings displayed the extreme variability in PPGR seen across individual participants, denoting the importance of personalized nutrition approaches in comparison to broad population-based recommendations[13].

Understanding food intake is a critical part of evaluating both population and personalized nutrition, and informatics approaches have also been used to track food purchasing and food intake and correlate it with nutrition information. One investigation examined whether grocery store purchases could be associated with specific nutrition information from a U.S. Department of Agriculture (USDA) database. The study found that most food products could be accurately mapped to nutritional composition[40]. While this investigation faced barriers mapping inconsistencies across food categories that have highly variant nutritional content, 70% of food items were mappable to the USDA nutrient database and 100% of items were mappable to USDA standard food groups. The investigators described a feasible approach for interpreting nutritional intake of grocery store purchases and expressed that greater interoperability between nutrition information and food labeling and production systems as well as healthcare would support translation of this type of research[40].

Utilizing electronic food diaries and phone applications has become a popular approach to documenting and analyzing dietary intake. Generally, electronic food diaries offer a wide range of functionality and the ability to store and share data across users, with aggregation and summarization of food intake and inferences on health outcomes offering users the most benefit[41]. Some versions of electronic food diaries within mobile device applications may include users sharing pictures of food they have consumed. With these types of applications, approximately a quarter of them provide either professional or crowdsourced feedback from other users[42]. Notably, most photo based apps are conducted with little to no application of evidence based methods for self-regulation and behavior change, which may impact user health behaviors and outcomes[42]. Beyond crowdsourced and professional responses to photos, one group of researchers developed an image recognition algorithm to recognize and analyze nutrition content from a photo of food[43]. This dietary tracking system, DietLens, utilizes deep-based food recognition technologies to classify the image and applies neural networks for image-level food categorization[43]. Thus far, this technology has been able to categorize food images from the research testing laboratory with between 75-99% accuracy, although difficulties were seen with mixed dishes that contain a large variation in ingredient composition[43]. When compared to other electronic food diary apps, DietLens displayed greater accuracy and required less time to log nutritional intake, indicating photo recognition based applications may be a useful tool for personal dietary intake tracking[43].

Overall, current progress in nutri-informatics research is promising and has given rise to novel findings and methodologies that can likely be utilized in future research endeavors. However, many barriers are still limiting the ability to bring nutri-informatics to the forefront of precision medicine and personalized health.

Current Nutri-informatics Challenges

Many clinical settings have focused on transitioning to electronic resources for nutritional data documentation and storage, allowing for widely accessible, albeit static clinical measurements. However, the multitude of methods for capturing nutritional information within EHRs and the lack of standardization across them limits their research use[39]. Furthermore, nutrition data is often sparse within EHRs, limiting the capacity to evaluate potential nutritional impacts on health outcomes[39]. As such, clinical nutri-informatics investigations more often focus on specific health outcomes such as specific disease states and clinical biomarkers as opposed to larger, more integrative studies that incorporate a wider array of data types and outcomes. A good example is the implementation of malnutrition screening assessments, which although they have been important for identifying and managing malnutrition, a validated Malnutrition Screening Tool (MST) asks only two questions, “Have you lost weight recently without trying? If yes, how much weight have you lost?" and "Have you been eating poorly because of a decreased appetite?"[44]. With just two questions asked via patient survey, no dietary intake or other information is acquired. This allows for the clinician to identify malnutrition at a gross level but provides little insight into any specific dietary factors. Very few attempts have been made to try and utilize nutrition and clinical survey data, -omics data, and other heterogeneous data types in coordination together in research.

Furthermore, major challenges exist for personalized nutrition endeavors, including experimental designs being unable to track the complex physiological response to nutrient exposures (nutrients and potential contaminants, food additives, or toxins), incomplete understanding and establishment of metabolic biomarkers, and inconsistent documentation language or incorrect reporting of health exposures and outcomes[39]. Concerns with self-report survey approaches also occur, as participants may inaccurately depict their nutritional intake, challenging research findings. In one investigation, average 24-hour dietary recalls underestimated dietary sodium intake, when compared to estimated consumption calculated from 24-hour urine sodium content[45]. Concerns for racial disparities utilizing food frequency questionnaires (FFQ) also arise. One investigation identified significantly greater correlation between 24-hour dietary recall and FFQ for white women compared to Black women, challenging the ability to decipher eating habits and make dietary recommendations given the racial disparities not captured within nutrition surveys[46].

Nutrition research generally faces challenges, with a variety of research methodologies available and each with tradeoffs of benefits and challenges. While single dietary element investigations are optimal for evaluating nutrition biomarkers, investigations that single out a particular dietary component can be challenging to complete in a controlled manner in humans and may also lead to broad interpretations regarding the systemic effects of the food or nutrient in health. On the opposing side, many investigations evaluate diet in its entirety, which limits the capacity for evaluation of specific cellular and molecular interactions or signaling[39]. For lack of a "perfect" methodological design, many nutrition investigations are conducted on similar concepts producing a wide range of data and data types that are not compatible enough for larger-scale insights, stunting opportunities for translational research and further hypothesis development.

Biomedical data as a whole is represented using a wide array of terminologies for similar or identical concepts, leading to challenges for data aggregation and management even with smaller data sets[47]. The National Institute of Health (NIH) has initiated the usage of common data elements (CDEs), which are data elements frequently assessed within research such as demographics, labs, or biomarkers that have been standardized using human and machine-readable definitions to facilitate research interoperability and consistency (https://cde.nlm.nih.gov/). CDEs are designed to support data collection and analysis in a consistent fashion[47] for a variety of data types such as from surveys, clinical data, and laboratory findings. CDEs are intended to be reused within and across projects, meaning two different assessments can ask the same question using the same CDE format and the data produced from the question will be compatible between surveys or studies. An example of a CDE used in the NHANES 1999-2000 questionnaire is depicted in Figure 2. (https://wwwn.cdc.gov/Nchs/Nhanes/1999-2000/ALQ.htm).

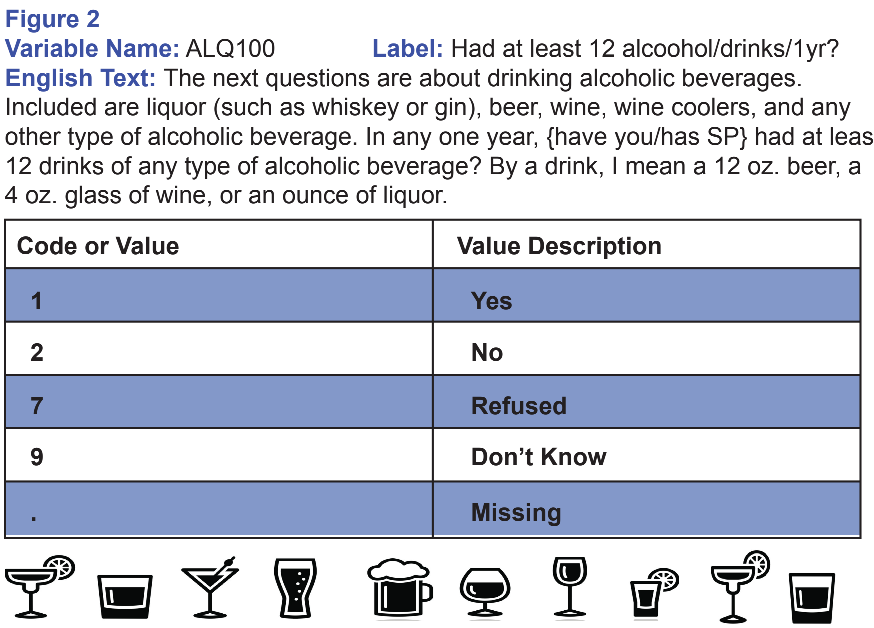


Figure 2

Variable Name: ALQ100 **Label:** Had at least 12 alcoohol/drinks/1yr?

English Text: The next questions are about drinking alcoholic beverages. Included are liquor (such as whiskey or gin), beer, wine, wine coolers, and any other type of alcoholic beverage. In any one year, {have you/has SP} had at least 12 drinks of any type of alcoholic beverage? By a drink, I mean a 12 oz. beer, a 4 oz. glass of wine, or an ounce of liquor.

Code or Value	Value Description
1	Yes
2	No
7	Refused
9	Don’t Know
.	Missing

Figure 2. Sample Common Data Element. This question is an example of a common data element (CDE) from the National Health and Nutrition Examination Survey (NHANES) 1999-2000 questionnaire. In this instance, survey participants were inquired about alcohol consumption throughout the year and their responses were standardized using the corresponding code/value pair. Usage of this CDE in a separate survey, such as a future year of NHANES, will allow data from both surveys to be directly comparable.

Utilizing CDEs, nutrition research and data could be approached with standardization and interoperability in mind, similarly to the structure CDEs have provided to projects within the fields of cancer research[47] and stroke clinical and epidemiological research[48]. CDEs and other scientific records like data sets can be managed via unique persistent identifiers (PIDs), which facilitate data sharing, reuse, and attribution[49]. While CDEs and PIDs for data reuse and sharing are commonly used in other areas of biomedical sciences, CDEs are not in widespread use in nutrition investigations. Furthermore, researchers have seldom discussed nutri-informatics research from the perspective of data reuse to maximize understanding and comparability of findings[50].

Biomedical ontologies can support standardization and integration of nutrition data

One approach to develop structure and standardization is the use of ontologies. Ontologies are classifications of terms focused on specific areas of knowledge or domains that include logically defined relationships between the terms[51]. Ontologies offer not only human readable definitions of terms, but also computer readable definitions in the form of logical definitions or axioms. This allows for reasoning across the data and increased computability[51–54]. While ontologies are created specifically for a domain or subdomain, many ontologies are co-developed to be interoperable and compatible with one another making it easier to exhibit relationships between terms in different ontologies[51]. Ontologies have been applied extensively in areas such as genomics and phenomics, which has allowed for increased connections between patient genotypes and clinical phenotypes, facilitating individualized medicine and rare disease identification[53–55].

Prominent ontologies frequently used in biomedical research include the Gene Ontology (GO) describing gene functions and biological processes[56], and the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT) which is a clinical terminology for medical conditions and symptoms[57]. While there are a variety of analyses and applications for utilizing ontologies, two common approaches are similarity comparisons and enrichment analyses. An example of a semantic similarity comparison is the use of non-exact phenotype profile matching. Using patient profiles encoded with phenotype terms from the Human Phenotype Ontology, multiple profiles can be compared to identify similar and unique phenotypes between them. The application of semantic similarity algorithms over ontology encoded clinical phenotype data for “fuzzy” phenotype matching has supported diagnosis of rare disease patients[58,59].

Enrichment analyses are also common approaches to utilizing ontologies. For example, an investigation looking to evaluate changes in vitamin D and serotonin gene expression for individuals with irritable bowel syndrome (IBS) assessed gene transcripts from tissue biopsy samples from IBS+ and IBS- populations[60]. After identifying genetic features of interest via differential expression, enrichment utilizing GO highlighted the associated pathways and functions of the differentially expressed genes[60]. In this instance, investigators identified the most prevalent enrichment within the serotonergic pathway, which paired with real-time polymerase chain reactions may indicate that IBS patient-derived RNA has lower tryptophan hydroxylase-1 expression, which is a rate limiting step in serotonin synthesis[60].

While the use of ontologies in scientific research has grown over the past few decades, their use within the discipline of nutrition has lagged regardless of researchers exhibiting a need for nutrition data standardization through the application of ontologies[50]. Integration of nutrition into biomedical ontologies holds the potential to identify hundreds of nutrition-disease, nutrition-phenotype, and nutrition-genotype relationships.

The complexity and coverage differ greatly across nutrition subdisciplines, but many nutrition-related knowledge resources do exist including some which can be leveraged to better understand nutrition and human health in a computable manner. In Tables 1 and 2, a review of prominent existing resources is provided, including food and nutrition focused ontologies and related biomedical knowledge resources.

Table 1. A listing of prominent nutri-informatics ontologies

Chemical Entities of Biological Interest (ChEBI)[68] https://www.ebi.ac.uk/chebi/	Description: A dictionary of molecular entities focused on ‘small’ chemical compounds. Includes chemical dietary metabolites and nutrients.
	Example Term Name: L-ascorbic acid
	Example Term ID: http://purl.obolibrary.org/obo/CHEBI_29073
	Example Term Definition: The L-enantiomer of ascorbic acid and conjugate acid of L-ascorbate.
	Example Term synonyms: (5R)-5-[(1S)-1,2-dihydroxyethyl]-3,4-dihydroxyfuran-2(5H)-one L-threo-hex-2-enono-1,4-lactone

Ontology of Nutritional Studies (ONS)[64] https://github.com/enpadasi/Ontology-for-Nutritional-Studies	Description: A systematic ontology framework for nutritional studies. Includes Nutrition study design and diets.
	Example Term Name: Intervention Diet
	Example Term ID: http://www.enpadasi.eu/ontology/release/v1/ons/ONS_0000081
	Example Term Definition: The diet administered during an intervention study. It usually comprises the adoption of a certain nutritional intervention, intended as the prescription of consuming or not consuming certain food, and follows a precise study design. Intervention studies usually compare at least two subgroups of a population, one control group receiving a null nutritional intervention, and one or more test groups receiving the intervention
Gene Ontology (GO)[56] http://geneontology.org/	Description: A computational model of functions of genes including metabolism related biological processes.
	Example Term Name: diacylglycerol metabolic process
	Example Term ID: http://purl.obolibrary.org/obo/GO_0046339
	Example Term Definition: The chemical reactions and pathways involving diacylglycerol, a glyceride in which any two of the R groups (positions not specified) are acyl groups while the remaining R group can be either H or an alkyl group.
	Example Term synonyms: diacylglycerol metabolism diglyceride metabolism
Food Biomarker Ontology (FOBI)[69] https://github.com/pcastellanoescuder/FoodBiomarkerOntology	Description: A joint ontology including Food Ontology and Biomarker Ontology including food intake biomarkers.
	Example Term Name: 3,4-dihydroxyphenylacetic acid
	Example Term ID: http://purl.obolibrary.org/obo/FOBI_030377
Medical Actions Ontology (MAxO) https://github.com/monarch-initiative/MAxO	Description: A structured vocabulary for medical procedures, interventions, therapies, and treatment including medical nutrition therapy.
	Example Term Name: dietary intervention
	Example Term ID: http://purl.obolibrary.org/obo/MAXO_0000088

	Example Term Definition: Any alteration or treatment in an individual's diet with a planned goal, usually designed to improve the individual's overall health.
	Example Term synonyms: behavioral nutritional intervention diet
Environmental Conditions, Treatments, and Exposures Ontology (ECTO) https://github.com/EnvironmentOntology/environmental-exposure-ontology	Description: A structured vocabulary for environmental exposures and stimuli including food, nutrient, and diet exposures.
	Example Term Name: vitamin D exposure
	Example Term ID: http://purl.obolibrary.org/obo/ECTO_9000133
	Example Term Definition: An exposure to vitamin D.
	Example Term synonyms: exposure to vitamin D
Human Phenotype Ontology (HPO)[70] https://hpo.jax.org/app/	Description: A standardized vocabulary of phenotypic abnormalities encountered in human disease including nutrition related phenotypes.
	Example Term Name: Abnormality of amino acid metabolism
	Example Term ID: http://purl.obolibrary.org/obo/HP_0004337
	Example Term Definition: Abnormality of an amino acid metabolic process.
	Example Term synonyms: amino acid levels abnormal
Mondo Disease Ontology (Mondo) https://mondo.monarchinitiative.org/	Description: A harmonization of disease definitions including nutrition related diseases.
	Example Term Name: protein-energy malnutrition
	Example Term ID: http://purl.obolibrary.org/obo/MONDO_0001371
	Example Term Definition: A nutritional deficit that is caused by inadequate protein or calorie intake.
	Example Term synonyms: protein energy malnutrition
Food Ontology (FoodOn)[63] https://foodon.org/	Description: An ontology focused on categorization and processing of food including terminology for food, food components, and food management.
	Example Term Name: macaroni and cheese

	Example Term ID: http://purl.obolibrary.org/obo/FOODON_00002960
	Example Term Definition: Macaroni and cheese is a dish of English origin, consisting of cooked macaroni pasta and a cheese sauce, most commonly cheddar. It can also incorporate other ingredients, such as breadcrumbs, meat and vegetables.
	Example Term synonyms: mac n cheese macaroni cheese mac and cheese
Neuro Behavioral Ontology (NBO)[71] https://github.com/obo-behavior/behavior-ontology	Description: An ontology of human and animal behaviors and behavioral phenotypes including motivation and behaviors regarding food and beverage consumption.
	Example Term Name: polyphagia
	Example Term ID: http://purl.obolibrary.org/obo/NBO_0000546
	Example Term Definition: A pathological eating behavior characterized by an abnormally large intake of food by mouth, usually due to excessive hunger that is relatively prolonged.
	Example Term synonyms: hyperphagia
Systematized Nomenclature of Medicine -- Clinical Terms (SNOMED CT)[57] http://www.snomed.org/snomed-ct/why-snomed-ct	Description: Comprehensive clinical terminology including nutrition related clinical terminology.
	Example Term Name: Vitamin D overdose (disorder)
	Example Term ID: SCTID: 296953002
	Example Child classes: Accidental vitamin D overdose (disorder) Intentional vitamin D overdose (disorder) Vitamin D overdose of undetermined intent (disorder)
International Classifications of Disease (ICD) https://icd.who.int/en	Description: The global standard for diagnostic health information including coding and billing terminology for nutrition diagnostics.
	Example Term: Wernicke's encephalopathy
	Example ICD-10-CM Code E51.2
	Example Term Definition: Wernicke's encephalopathy (or Wernicke's disease) refers to the presence of neurological symptoms caused by biochemical lesions of the central nervous system after exhaustion of B-vitamin reserves, in particular thiamine (vitamin B1). The condition is part of a larger

	group of diseases related to thiamine insufficiency, including beriberi in all its forms, and Korsakoff syndrome. When Wernicke's encephalopathy occurs simultaneously with Korsakoff syndrome it is known as Wernicke–Korsakoff syndrome.
	Example Specialty: Endocrinology
Medical Subject Headings (MeSH)[72] https://www.nlm.nih.gov/mesh/meshhome.html	Description: A vocabulary of biomedical and health-related information including nutrients and food components.
	Example Term Name: Niacin
	Example Term ID: D009525
	Example Term Definition: A water-soluble vitamin of the B complex occurring in various animal and plant tissues. It is required by the body for the formation of coenzymes NAD and NADP. It has PELLAGRA-curative, vasodilating, and antilipemic properties.
	Example Entry names: 3-Pyridinecarboxylic Acid Niacin Aluminum Salt Niacin Ammonium Salt Niacin Calcium Salt

Table 2. A listing of prominent nutri-informatics databases and resources

Comparative Toxicogenomics Database[73] http://ctdbase.org/	Description: A publicly available dataset of environmental exposure impacts on human health including representation of food chemical and metabolite interactions and nutrition related disease, phenotype, and gene associations.
	Example Term Name: Folic Acid
	Example Term Definition: A member of the vitamin B family that stimulates the hematopoietic system. It is present in the liver and kidney and is found in mushrooms, spinach, yeast, green leaves, and grasses (POACEAE). Folic acid is used in the treatment and prevention of folate deficiencies and megaloblastic anemia.
	Example Top 3 Gene Interactions: SLC19A1, TNF, AGT
	Example Term Synonyms: B9, Vitamin | Folacin | Folate | Folic Acid, Calcium Salt (1:1) | Folic Acid, (D)-Isomer | Folic Acid, (DL)-Isomer | Folic Acid, Monopotassium Salt | Folic Acid, Monosodium Salt | Folic Acid, Potassium Salt | Folic Acid, Sodium

	Salt | Folvite | Pteroylglutamic Acid | Vitamin B9 | Vitamin M
NutriGenomeDB[74] http://nutrigenomedb.org/	Description: A nutrigenomics exploratory and analytical platform including nutrigenomics gene expression data modules.
	Primary Features: Gene Expression Browser – Includes gene expression database search tools and expression heat map generation tool. Phenotype-Centered Analysis – Evaluates a list of differentially expressed genes and aggregates similar experiments based on gene expression profiles characterizing specific phenotypes.
The Monarch Initiative[75] https://monarchinitiative.org/	Description: An integrative data and analytic platform connecting data across species including nutrition related phenotypes, genotypes, and diseases.
	Example Term Name: rickets (disease)
	Example Term ID: http://purl.obolibrary.org/obo/MONDO_0005520
	Example Term Definition: Bone softening and weakening usually caused by deficiency or impaired metabolism of vitamin D. Deficiency of calcium, magnesium, or phosphorus may also cause rickets. It predominantly affects children who suffer from severe malnutrition. It manifests with bone pain, fractures, muscle weakness, and skeletal deformities.
	Example Exact synonyms: vitamin D-dependent rickets, rachitis, vitamin D hydroxylation-deficient rickets, rickets Example Related Synonyms: active rickets, vitamin-D deficiency rickets, hypovitaminosis D, nutritional rickets, vitamin D deficiency disease
	Example Related Phenotypes: short stature, tibial bowing, recurrent fractures
	Example Genes (causal): PHEX, FGF23, SLC34A3
USDA FoodData Central (FDC) https://fdc.nal.usda.gov/index.html	Description: An integrated data system with nutrient profile data and links to related research including nutrient profiles for common foods and beverages.
	Example Term Name: Beans, black turtle, mature seeds, canned
	Example FDC ID: 175188

	Example Food Category: Legumes and Legume Products
	Available data includes: Macronutrient content Micronutrient content Average weight/volume
USDA National Health and Nutrition Examination Survey (NHANES) and What We Eat in America (WWEIA)[29] data sets https://www.cdc.gov/nchs/nhanes/index.htm	Description: Data sets of compiled health and nutrition status information for adults and children.
	NHANES available data: • Biological samples of serum, plasma, urine, and DNA • Demographic, Dietary, Examination, Laboratory, and Questionnaire health data WWEIA available data: 24-hour dietary recalls (2) • Type and amount of food and beverage consumed • Time, location, and name of eating occasion • Water consumption type and source • Use of table salt • Special diets and usual daily intake • Calculated daily total intake of energy and 60+ nutrients/food components

Nutri-informatics resources are largely focused on clinical nutrition, foods, and nutrients (Table 1). Terminological resources such as the Mondo Disease Ontology, the Human Phenotype Ontology (HPO), SNOMED-CT, and International Classification of Diseases (ICD) have nutrition related diseases and phenotypes. The Monarch Initiative and the Comparative Toxicogenomics Database (CTD) denote relationships between nutrients, disease, phenotypes, and genes. Because nutrition can impact diseases or phenotypes that do not have an exclusively nutrition-based etiology, it is imperative that such relationships are discovered and included in these resources. Gene pathways for metabolism are represented in GO and nutrient-gene expression analysis platforms such as NutriDB have been introduced. Also, initial modeling of nutrient exposures in the Environmental Conditions, Treatments, and Exposures Ontology (ECTO), dietary patterns and interventions in the Ontology for Nutritional Studies (ONS), and nutrient therapies in the Medical Action Ontology (MAxO) have been represented, alongside a few nutrition related behaviors in the Neuro Behavior Ontology (NBO). There is also a need for representation of public health nutrition investigations, and resources such as the Ontology for Nutritional Epidemiology (ONE)[61] are promising. Such knowledge representation is still emergent and will continue to grow in the years to come and require further nutrition representation.

Foods and nutrients have strong representation in the Food Ontology (FoodOn) and US Department of Agriculture Food Data Central (USDA FDC), and food processing is included in FoodOn. Macro- and micronutrients are well represented in the Chemical Entities of Biological Interest (ChEBI), Medical Subject Headings (MeSH), and CTD. Food related biomarkers and metabolites that can be identified in biological samples are seen in ChEBI and the Food-Biomarker Ontology (FOBI). However, the large multitude of foods and beverages available for consumption as well as the wide array of agricultural and processing techniques used with consumable products requires much more extensive representation. Furthermore, nutrition biomarkers are a developing field and these resources will require continual revision.

Representation is still greatly lacking in areas such as nutrition sociology (e.g. food behaviors, beliefs, culture, norms, nutrition literacy), public health nutrition policy, and nutrition education. While nutrition and food representation in ontologies and databases will require substantial work from the ontology and data science and nutrition research communities to ensure adequate representation, there are still meaningful relationships represented in current resources. Figure 3 depicts meaningful relationships between food, agriculture, phenotypes, and disease that can currently be represented using biomedical ontologies.

Figure 3A

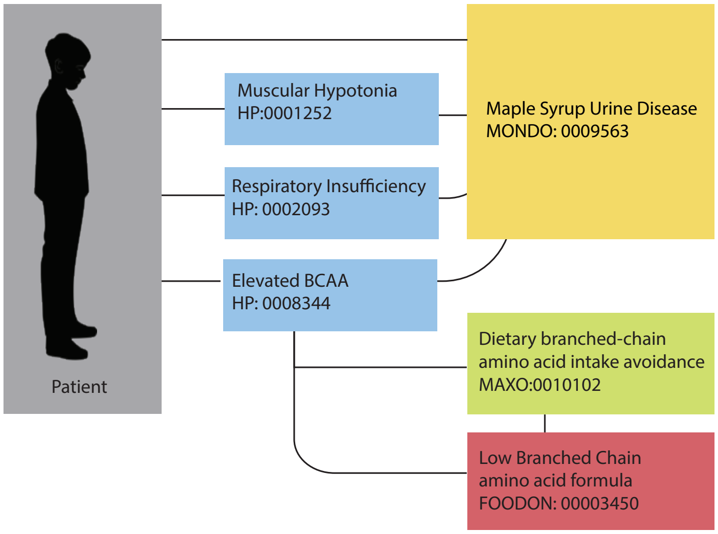


Figure 3B

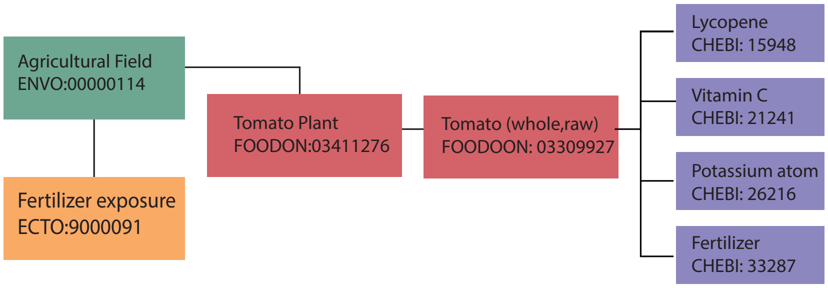


Figure 3. Representing nutrition using ontologies. Nutrition representation in current ontologies and databases is not yet sufficient to meet the needs of the nutri-informatics research community, yet some meaningful relationships can still be identified within the current landscape. **3A. Maple Syrup Urine Disease**. This rare metabolic disease can be annotated with related phenotypes, genes, nutritional recommendations, and medical foods using interoperable biomedical ontology terms. **3B. Farm to Fork with a Tomato.** The process of growing a tomato can also be annotated by its exposures and nutrient content.

Opportunities for integrative nutri-informatics research

As organizations such as the NIH establish priorities for investigating nutrition exposures and precision nutrition for human health[62], the field of nutrition is in a meaningful position to launch investigations and continue growing our understanding of clinical dietary solutions for individuals and populations via nutri-informatics. In order to continue the integration of nutrition data into existing knowledge graphs and ontologies, comprehensive, standardized representation of all categories of nutrition from basic science to public policy is needed. This requires progressive integration of the many essential categories of nutrition seen in Figure 1, including foods and nutrients, clinical disease management through nutrition as well as those currently underrepresented categories such as sociological impacts on nutrition. In order to achieve this representation, there are multiple hurdles and opportunities that need to be addressed by the biomedical ontology and nutrition research communities:

1) **Incomplete coverage of nutrition-related concepts in ontologies.** Due to the widespread field of nutrition, representation for all subdisciplines of nutrition has yet to be achieved. Although efforts such as Food Ontology (FoodOn)[63] and Ontology of Nutritional Studies (ONS)[64] as well as others have made strides in representing foods and food components as well as nutrition intervention and epidemiological terminology, further areas such as nutrition sociology, nutrition policies, and nutrition education are still limited.
2) **Creation of new relationships and modeling of nutrition as a factor in disease and phenotype presentation, prevention, and management is needed.** While existing knowledge bases such as The Monarch Initiative[65] may address nutrition concepts like disease states from nutrient deficiencies, the focus is on other biomedical fields such as genomics and the impact on disease and phenotypes. It is likely that nutrition-related diseases and nutrition impacts on disease are underrepresented currently.
3) **Limited compatibility across databases and knowledge resources containing nutrition related information due to a lack of community development standards.** Community standards for ontology and knowledge base development are not established for many nutrition resources, limiting their compatibility with other nutrition- and biomedically-focused computable resources. Without a set of standards for data management and interaction between researchers across the community, nutrition data will not be fully integrable.
4) **Poor communication and accountability regarding the FAIR principles of scientific data management and stewardship amongst nutrition researchers, nutrition journals, and nutrition research funding agencies.** Findability, Accessibility, Interoperability, and Reusability (FAIR) have been designated as the foundational principles to guide data production and publication to support data transparency and maximize data outcomes[66]. Due to the limited requirements or even recommendations for nutrition researchers to adhere to FAIR principles during their experimental process, scientific journals and research funding agencies limit the production and publication of optimal, computable data. This not only limits nutrition research findings, but also hampers knowledge gains in related biomedical fields.

Given the lack of FAIR principles utilized and required for current nutrition research, and the need to integrate nutrition data and knowledge into existing knowledge bases, there is a need for standardized nutrition vocabulary, as well as best practices to encourage the curation of nutrition related phenotypes, nutrition exposures, nutrition sociology, and other nutrition subfields. This is especially true as investigations into environmental exposures, including nutrition, continue to be pursued in relation to the genome and gene expression[67].

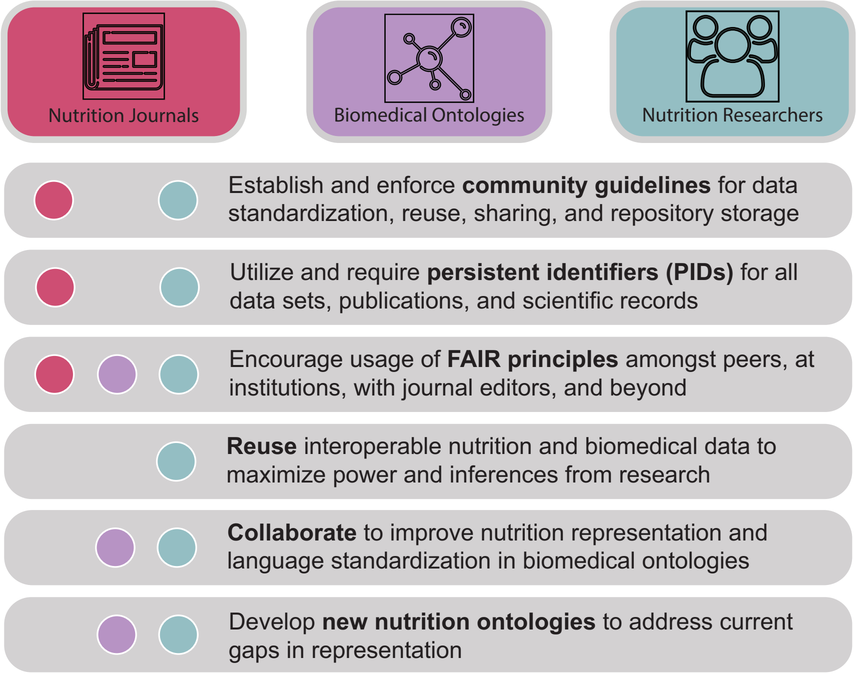
A Call for Improved Nutrition Representation and Standards

In order to further our understanding of how nutrition and food impacts health, human behavior, culture, and beyond, integrating nutrition terminology and relationships into ontologies and knowledge resources is essential. Increasing representation of nutrition in areas already being modeled in ontologies, such as foods, should be a focus, as well as areas yet to be explored in current resources. Areas including nutrition biomarkers, nutrition behavioral counseling, nutritional personal and cultural beliefs, and food processing can fall into this category.

The resources in Tables 1 and 2 have started striving towards representing nutrition in some capacity, but due to the vast nature of the field of nutrition, this representation is still incomplete for many topics as they have yet to be developed. Furthermore, some nutrition resources are not developed with compatibility in mind, further limiting interpretation and alignment of terminology across resources. These challenges in nutrition representation and compatibility across resources will require substantial, consistent work from the nutrition and ontology communities.

Working towards this goal of nutrition representation, a working group including curators from MAxO, ECTO, FoodOn, ONS, FOBI, the US Department of Agriculture, and other representatives are meeting regularly to discuss nutrition in the current ontology landscape. Thus far, this group has focused on how to represent diet, how to model an organism's biological capacity to consume certain foods, and the agricultural production related to foods. This working group functions via a GitHub page (https://github.com/FoodOntology/joint-food-ontology-wg) and is open to individuals or groups interested in participating.

Beyond working groups such as this, further steps towards nutrition representation in this landscape are needed from nutrition researchers, academic nutrition journals and publishers, and biomedical ontology developers and curators, which are described in Figure 4.



Nutrition Journals	Biomedical Ontologies	Nutrition Researchers	
●		●	Establish and enforce **community guidelines** for data standardization, reuse, sharing, and repository storage
●		●	Utilize and require **persistent identifiers (PIDs)** for all data sets, publications, and scientific records
●	●	●	Encourage usage of **FAIR principles** amongst peers, at institutions, with journal editors, and beyond
		●	**Reuse** interoperable nutrition and biomedical data to maximize power and inferences from research
	●	●	**Collaborate** to improve nutrition representation and language standardization in biomedical ontologies
	●	●	Develop **new nutrition ontologies** to address current gaps in representation

Figure 4. A Call to Action. Nutri-informatics stakeholders such as nutrition researchers, biomedical ontology developers, and academic journal communities are needed to realize the connectivity and analyzability of nutrition data. Key tasks are described here, including actions to improve data interoperability, identifiability, and collaboration between communities.

With greater understanding of the critical need for data standards in nutrition and the subsequent enforcement of those standards, nutrition researchers, journals, and ontology curators can maximize the research outcomes in nutri-informatics and related biomedical fields, supporting data interoperability and reuse in biomedical sciences. By representing nutrition semantics within biomedical ontologies, all currently represented biological fields can be correlated with nutrition and dietary exposures, including connections to diseases, phenotypes, and genes. Beyond the clinical realm, representation of cultural food, agronomical practices, personal beliefs regarding diet, public health nutrition policies, and the many other subdisciplines of nutrition all hold substantial potential for computable research in nutri-informatics.

With the utilization of biomedical ontologies and development of nutrition community standards for supporting FAIR principles, nutri-informatics research can progress to develop similar investigations to that of other fields. Nutrition data within ontologies may offer the ability to evaluate the impacts of dietary patterns, food combinations, pesticides and agricultural chemical exposures, cultural values, and individual behavioral impacts on human health. While these nutri-informatics investigations may not be achievable with the current nutrition ontology resources, further development in this field will undoubtedly offer novel understandings of how nutrition impacts human life.

Conclusion

Nutrition is a fundamental component to human and non-human animal life as an integral factor in the presentation of diseases, genes, or phenotypes, as well as an influencing factor on behavior and culture. While modern nutrition research may be a “younger” field of biology, it is far from insignificant and will require robust community standards in order to fully support FAIR practices, transparency, and maximal knowledge gains from research. By utilizing standardized language and biomedical ontologies, nutrition data could be integrated into the larger scheme of biomedical knowledge bases, supporting interoperability and reuse. This integration work is already beginning with the initiation of new nutrition focused ontologies and working groups. Through continued education and action from nutrition researchers and ontology developers to integrate nutrition research into biomedical ontologies, nutri-informatics investigations can grow to their full potential, supporting discovery from nutrition data beyond a single investigation and offering insights beyond the field of nutrition.

Acknowledgements:

This work was supported by a NIH Office of the Director Grant #5R24OD011883.

References:

[1] Mozaffarian D, Rosenberg I, Uauy R. History of modern nutrition science—implications for current research, dietary guidelines, and food policy. The BMJ 2018;361.

[2] Jew S, Antoine J-M, Bourlioux P, et al. Nutrient essentiality revisited. J Funct Foods 2015;14:203–9.

[3] Grimble GK. Essential and Conditionally-Essential Nutrients in Clinical Nutrition. Nutr Res Rev 1993;6:97–119.

[4] Chipponi J, Bleier J, Santi M, et al. Deficiencies of essential and conditionally essential nutrients. Am J Clin Nutr n.d.

[5] Schmid A, Walther B. Natural Vitamin D Content in Animal Products. Adv Nutr 2013;4:453–62.

[6] Lazarte CE, Carlsson N-G, Almgren A, et al. Phytate, zinc, iron and calcium content of common Bolivian food, and implications for mineral bioavailability. J Food Compos Anal 2015;39:111–9.

[7] Traber MG, Packer L. Vitamin E: beyond antioxidant function. Am J Clin Nutr 1995;62:1501S-1509S.

[8] Chung HR. Iodine and thyroid function. Ann Pediatr Endocrinol Metab 2014;19:8–12.

[9] Fennema O. Effects of Freeze Preservation on Nutrients. In: Karmas E, Harris RS, editors. Nutr. Eval. Food Process., Dordrecht: Springer Netherlands; 1988, p. 269–317.

[10] Hotz C, Gibson RS. Traditional Food-Processing and Preparation Practices to Enhance the Bioavailability of Micronutrients in Plant-Based Diets. J Nutr 2007;137:1097–100.

[11] Reddy MB, Love M. The Impact of Food Processing on the Nutritional Quality of Vitamins and Minerals. In: Jackson LS, Knize MG, Morgan JN, editors. Impact Process. Food Saf., Boston, MA: Springer US; 1999, p. 99–106.

[12] Kussmann M, Fay LB. Nutrigenomics and personalized nutrition: science and concept. Pers Med 2008;5:447–55.

[13] Zeevi D, Korem T, Zmora N, et al. Personalized Nutrition by Prediction of Glycemic Responses. Cell 2015;163:1079–94.

[14] Bush CL, Blumberg JB, El-Sohemy A, et al. Toward the Definition of Personalized Nutrition: A Proposal by The American Nutrition Association. J Am Coll Nutr 2020;39:5–15.

[15] Challa HJ, Ameer MA, Uppaluri KR. DASH Diet (Dietary Approaches to Stop Hypertension). StatPearls, Treasure Island (FL): StatPearls Publishing; 2020.

[16] Tosti V, Bertozzi B, Fontana L. Health Benefits of the Mediterranean Diet: Metabolic and Molecular Mechanisms. J Gerontol Ser A 2018;73:318–26.

[17] Rosanoff A, Dai Q, Shapses SA. Essential Nutrient Interactions: Does Low or Suboptimal Magnesium Status Interact with Vitamin D and/or Calcium Status? Adv Nutr 2016;7:25–43.

[18] Combet E, Gray SR. Nutrient–nutrient interactions: competition, bioavailability, mechanism and function in health and diseases. Proc Nutr Soc 2019;78:1–3.

[19] Dahan A, Altman H. Food–drug interaction: grapefruit juice augments drug bioavailability — mechanism, extent and relevance. Eur J Clin Nutr 2004;58:1–9.

[20] Xu K-Y, Xia G-H, Lu J-Q, et al. Impaired renal function and dysbiosis of gut microbiota contribute to increased trimethylamine-N-oxide in chronic kidney disease patients. Sci Rep 2017;7:1445.

[21] Basolo A, Hohenadel M, Ang QY, et al. Effects of underfeeding and oral vancomycin on gut microbiome and nutrient absorption in humans. Nat Med 2020;26:589–98.

[22] Fieldhouse P. Food and Nutrition: Customs and culture. Springer; 2013.

[23] Olson EL. The rationalization and persistence of organic food beliefs in the face of contrary evidence. J Clean Prod 2017;140:1007–13.

[24] Nikolaus CJ, Nickols-Richardson SM, Ellison B. Wasted food: A qualitative study of U.S. young adults' perceptions, beliefs and behaviors. Appetite 2018;130:70–8.

[25] Boatemaa S, Badasu DM, de-Graft Aikins A. Food beliefs and practices in urban poor communities in Accra: implications for health interventions. BMC Public Health 2018;18:434.

[26] Coveney J, O'Dwyer LA. Effects of mobility and location on food access. Health Place 2009;15:45–55.

[27] Knowles M, Rabinowich J, Ettinger de Cuba S, et al. "Do You Wanna Breathe or Eat?": Parent Perspectives on Child Health Consequences of Food Insecurity, Trade-Offs, and Toxic Stress. Matern Child Health J 2016;20:25–32.

[28] Döring F, Rimbach G. Nutri-informatics: a new kid on the block? Genes Nutr 2014;9:1–3.

[29] Group FSR, Health USDO, Control CFD, et al. What We Eat In America (WWEIA) 2015.

[30] Cifelli C, Auestad N, Fulgoni III V. Replacing the nutrients in dairy foods with non-dairy foods will increase cost, energy intake and require large amounts of food: National Health and Nutrition Examination Survey 2011–2014. Public Health Nutr 2020:1–12.

[31] Sattgast LH, Gallo S, Frankenfeld CL, et al. Nutritional Intake and Bone Health Among Adults With Probable Undiagnosed, Untreated Celiac Disease: What We Eat in America and NHANES 2009–2014. J Am Coll Nutr 2020;39:112–21.

[32] Hu FB, Satija A, Rimm EB, et al. Diet Assessment Methods in the Nurses' Health Studies and Contribution to Evidence-Based Nutritional Policies and Guidelines. Am J Public Health 2016;106:1567–72.

[33] Petimar J, Smith-Warner SA, Fung TT, et al. Recommendation-based dietary indexes and risk of colorectal cancer in the Nurses' Health Study and Health Professionals Follow-up Study. Am J Clin Nutr 2018;108:1092–103.

[34] Rusnak S, Charney P. Position of the Academy of Nutrition and Dietetics: Nutrition Informatics. J Acad Nutr Diet 2019;119:1375–82.

[35] Molinar LS, Childers AF, Hoggle L, et al. Informatics Initiatives at the Academy of Nutrition and Dietetics. J Acad Nutr Diet 2017;117:1293–301.

[36] Rossi M, Campbell KL, Ferguson M. Implementation of the Nutrition Care Process and International Dietetics and Nutrition Terminology in a Single-Center Hemodialysis Unit: Comparing Paper vs Electronic Records. J Acad Nutr Diet 2014;114:124–30.

[37] Trtovac D, Lee J. The Use of Technology in Identifying Hospital Malnutrition: Scoping Review. JMIR Med Inform 2018;6:e4.

[38] Eglseer D, Halfens RJG, Lohrmann C. Use of an electronic malnutrition screening tool in a hospital setting: effects on knowledge, attitudes and perceived practices of healthcare staff. Br J Nutr 2018;120:150–7.

[39] Verma M, Hontecillas R, Tubau-Juni N, et al. Challenges in Personalized Nutrition and Health. Front Nutr 2018;5.

[40] Brinkerhoff KM, Brewster PJ, Clark EB, et al. Linking Supermarket Sales Data To Nutritional Information: An Informatics Feasibility Study. AMIA Annu Symp Proc 2011;2011:598–606.

[41] Rusin M, Årsand E, Hartvigsen G. Functionalities and input methods for recording food intake: A systematic review. Int J Med Inf 2013;82:653–64.

[42] Hales S, Dunn C, Wilcox S, et al. Is a Picture Worth a Thousand Words? Few Evidence-Based Features of Dietary Interventions Included in Photo Diet Tracking Mobile Apps for Weight Loss. J Diabetes Sci Technol 2016;10:1399–405.

[43] Ming Z-Y, Chen J, Cao Y, et al. Food Photo Recognition for Dietary Tracking: System and Experiment. In: Schoeffmann K, Chalidabhongse TH, Ngo CW, et al., editors. Multimed. Model., Cham: Springer International Publishing; 2018, p. 129–41.

[44] Ferguson M, Capra S, Bauer J, et al. Development of a valid and reliable malnutrition screening tool for adult acute hospital patients. Nutrition 1999;15:458–64.

[45] McLean R, Cameron C, Butcher E, et al. Comparison of 24-hour urine and 24-hour diet recall for estimating dietary sodium intake in populations: A systematic review and meta-analysis. J Clin Hypertens 2019;21:1753–62.

[46] Olendzki B, Procter-Gray E, Magee MF, et al. Racial differences in misclassification of healthy eating based on food frequency questionnaire and 24-hour dietary recalls. J Nutr Health Aging 2017;21:787–98.
[47] Nadkarni PM, Brandt CA. The Common Data Elements for Cancer Research: Remarks on Functions and Structure. Methods Inf Med 2006;45:594–601.
[48] Saver Jeffrey L., Warach Steven, Janis Scott, et al. Standardizing the Structure of Stroke Clinical and Epidemiologic Research Data. Stroke 2012;43:967–73.
[49] Dappert A, Farquhar A, Kotarski R, et al. Connecting the Persistent Identifier Ecosystem: Building the Technical and Human Infrastructure for Open Research. Data Sci J 2017;16:28.
[50] Lemay DG, Zivkovic AM, German JB. Building the bridges to bioinformatics in nutrition research. Am J Clin Nutr 2007;86:1261–9.
[51] Washington N, Lewis S. Ontologies: Scientific Data Sharing Made Easy. Nat Educ 2008;1:5.
[52] Rubin DL, Lewis SE, Mungall CJ, et al. National Center for Biomedical Ontology: Advancing Biomedicine through Structured Organization of Scientific Knowledge. OMICS J Integr Biol 2006;10:185–98.
[53] Mabee P, Ashburner M, Cronk Q, et al. Phenotype ontologies: the bridge between genomics and evolution. Trends Ecol Evol 2007;22:345–50.
[54] Haendel MA, Chute CG, Robinson PN. Classification, Ontology, and Precision Medicine. N Engl J Med 2018;379:1452–62.
[55] Westbury SK, Turro E, Greene D, et al. Human phenotype ontology annotation and cluster analysis to unravel genetic defects in 707 cases with unexplained bleeding and platelet disorders. Genome Med 2015;7:36.
[56] The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res 2019;47:D330–8.
[57] El-Sappagh S, Franda F, Ali F, et al. SNOMED CT standard ontology based on the ontology for general medical science. BMC Med Inform Decis Mak 2018;18:76.
[58] Masino AJ, Dechene ET, Dulik MC, et al. Clinical phenotype-based gene prioritization: an initial study using semantic similarity and the human phenotype ontology. BMC Bioinformatics 2014;15:248.
[59] Köhler S, Schulz MH, Krawitz P, et al. Clinical Diagnostics in Human Genetics with Semantic Similarity Searches in Ontologies. Am J Hum Genet 2009;85:457–64.
[60] Dussik CM, Hockley M, Grozić A, et al. Gene Expression Profiling and Assessment of Vitamin D and Serotonin Pathway Variations in Patients With Irritable Bowel Syndrome. J Neurogastroenterol Motil 2018;24:96–106.
[61] Yang C, Ambayo H, De Baets B, et al. An Ontology to Standardize Research Output of Nutritional Epidemiology: From Paper-Based Standards to Linked Content. Nutrients 2019;11.
[62] Rodgers GP, Collins FS. Precision Nutrition—the Answer to “What to Eat to Stay Healthy.” JAMA 2020.
[63] Dooley DM, Griffiths EJ, Gosal GS, et al. FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration. Npj Sci Food 2018;2:1–10.
[64] Vitali F, Lombardo R, Rivero D, et al. ONS: an ontology for a standardized description of interventions and observational studies in nutrition. Genes Nutr 2018;13:12.
[65] Mungall CJ, McMurry JA, Köhler S, et al. The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species. Nucleic Acids Res 2017;45:D712–22.
[66] Wilkinson MD, Dumontier M, Aalbersberg IjJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016;3:160018.
[67] Patel CJ, Pho N, McDuffie M, et al. A database of human exposomes and phenomes from the US National Health and Nutrition Examination Survey. Sci Data 2016;3:1–10.

[68] Hastings J, Owen G, Dekker A, et al. ChEBI in 2016: Improved services and an expanding collection of metabolites. Nucleic Acids Res 2016;44:D1214–9.
[69] Castellano-Escuder P, González-Domínguez R, Wishart DS, et al. FOBI: an ontology to represent food intake data and associate it with metabolomic data. Database 2020;2020.
[70] Köhler S, Carmody L, Vasilevsky N, et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. Nucleic Acids Res 2019;47:D1018–27.
[71] Gkoutos GV, Schofield PN, Hoehndorf R. The Neurobehavior Ontology: An Ontology for Annotation and Integration of Behavior and Behavioral Phenotypes. In: Chesler EJ, Haendel MA, editors. Int. Rev. Neurobiol., vol. 103, Academic Press; 2012, p. 69–87.
[72] Lipscomb CE. Medical Subject Headings (MeSH). Bull Med Libr Assoc 2000;88:265–6.
[73] Davis AP, Grondin CJ, Johnson RJ, et al. The Comparative Toxicogenomics Database: update 2019. Nucleic Acids Res 2019;47:D948–54.
[74] Martín-Hernández R, Reglero G, Ordovás JM, et al. NutriGenomeDB: a nutrigenomics exploratory and analytical platform. Database 2019;2019.
[75] Shefchek KA, Harris NL, Gargano M, et al. The Monarch Initiative in 2019: an integrative data and analytic platform connecting phenotypes to genotypes across species. Nucleic Acids Res 2020;48:D704–15.