

APPROVED: 31 January 2022

doi: 10.2903/j.efsa.2022.e200411

Appraising diet–disease associations to be used in risk assessment, including an insight in nutritional epidemiology

Vânia Mendes, Aikaterini Niforou and Androniki Naska

Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Greece

Abstract

High consumption of red meat, in particular of its processed products, has been linked to the development of various chronic diseases, and the need to reduce consumption levels of these products has been identified as a public health priority in Europe. Among the potential alternatives, pulses have gained a prominent position in recent years. Stemming from the broader context of the substitution of red meat in the diet with alternative plant-based protein sources, this work programme was developed to address the need of improving the current understanding on the public health impact of potential substitutes. From a training perspective, the main goal was to advance the fellow's background knowledge in the principles of nutritional epidemiology, while contributing to the harmonisation of food risk assessment practices across Europe. The activities developed comprised a comprehensive learning experience in the different components of the risk assessment framework, giving particular focus to the appraisal of epidemiological evidence within this context (hazard identification and characterisation). The fellow had the opportunity to gain hands-on experience with EFSA's databases among other relevant national and international data sources. The association between the consumption of legumes and risk of cardiovascular disease in particular was extensively explored and described by summarising the available evidence through dose–response meta-analyses. The one-stage approach method was used to explore the shape of the associations, including studies with as few as two levels of exposure, in a weighted mixed-effects model. The analysis was also performed taking into account the potential risk of bias of included studies, assessed using the ROBINS-I tool. Upon the completion of the activities carried out, the fellow made significant progress towards the learning outcomes expected to be achieved throughout the duration of the EU-FORA fellowship period.

© 2022 Wiley-VCH Verlag GmbH & Co. KGaA on behalf of the European Food Safety Authority.

Keywords: risk assessment, plant foods, legumes, pulses, dose–response association, risk of bias

Correspondence: eu-fora@efsa.europa.eu

Declarations of interest: The declarations of interest of all scientific experts active in EFSA's work are available at <https://ess.efsa.europa.eu/doi/doiweb/doisearch>.

Acknowledgements: This report is funded by EFSA as part of the EU-FORA programme. The authors would like to thank Maria-Iosifina Kasdagli, biostatistician and Ermolaos Ververis, PhD candidate at the Dept. of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens and Scientific Officer at EFSA NUTRI-UNIT for their scientific support throughout the programme.

Suggested citation: Mendes V, Niforou A and Naska A, 2022. Appraising diet–disease associations to be used in risk assessment, including an insight in nutritional epidemiology. EFSA Journal 2022;20(S1): e200411, 15 pp. <https://doi.org/10.2903/j.efsa.2022.e200411>

ISSN: 1831-4732

© 2022 Wiley-VCH Verlag GmbH & Co. KGaA on behalf of the European Food Safety Authority.

This is an open access article under the terms of the [Creative Commons Attribution-NoDerivs](https://creativecommons.org/licenses/by-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited and no modifications or adaptations are made.



The EFSA Journal is a publication of the European Food Safety Authority, a European agency funded by the European Union.



Table of contents

Abstract.....	1
1. Introduction.....	4
2. Description of work programme	4
2.1. Aims.....	4
2.2. Activities/Methods	4
2.2.1. Identification of the subject matter and patterns of consumption of pulses in Greece and beyond.....	4
2.2.1.1. Definition of the exposure.....	5
2.2.1.2. Assessment of trends in the consumption of pulses worldwide, in Europe and Greece	5
2.2.2. Understanding of the current state of the art: health outcomes associated with the consumption of legumes	6
2.2.3. Assessment of dose-response associations between the consumption of legumes and the selected health outcome.....	8
2.2.3.1. Literature search and selection of studies.....	8
2.2.3.2. Data extraction and dose calculation.....	9
2.2.3.3. Risk of bias assessment	9
2.2.3.4. Dose-response analysis.....	10
2.2.4. Additional activities.....	10
2.2.4.1. Preliminary literature review on anti-nutrients present in pulses	10
3. Conclusions.....	11
References.....	11
Abbreviations.....	12
Appendix A – Comparison between EFSA FoodEx21 and other food classification systems for the food group of species in the <i>Leguminosae</i> family	13
Appendix B – Additional activities.....	15

1. Introduction

The consumption of red meat has significantly increased during the last five decades. Despite being an important source of several macro- and micronutrients, high levels of red meat in the diet, especially of its processed products, have been linked with various chronic diseases. Hence, a reduction in the consumption of red meat has been identified as a public health priority in Europe. Reasons such as consumers' increasing concern and preference for health promoting foods have also led to the emergence of products that can replace red meat. Among the potential alternatives, pulses have gained a prominent position. Cultivation and consumption of pulses have a long tradition in almost all regions of the world and, for centuries, this food group has played a fundamental role in the functioning of traditional agricultural systems and dietary patterns of populations. Pulses are extremely nutritious, with high levels of protein, dietary fibre and with a very low fat content. Moreover, they offer agricultural and ecological benefits in terms of nitrogen fixation in the soil and have a low water and carbon footprint. Despite their nutritional and environmental benefits, in the recent decades, pulses have not received deserved attention as an important food component in the diet of populations worldwide. More recently, higher concerns with health, nutrition and the environment seem to be leading to significant changes in food consumption patterns and driving a shift to more plant-based inclusive diets. This work programme was developed within the broader context of the substitution of red meat in the diet with alternative plant-based protein sources, addressing the need to better understand the impact of potential substitutes on public health. This report describes the knowledge and experience acquired by the fellow during the EU-FORA fellowship programme based at the Dept. of Hygiene, Epidemiology and Medical Statistics of the School of Medicine, National and Kapodistrian University of Athens, as well as a synthesis of results achieved.

2. Description of work programme

2.1. Aims

Using the *Leguminosae* family species as a case study, the programme's main goal was to advance the fellow's background knowledge in the principles of nutritional epidemiology, while contributing to the harmonisation of food risk assessment practices across Europe, namely those related to the appraisal of epidemiological evidence, which are inextricably related to hazard identification and characterisation. The specific learning objectives for this work programme were (1) to increase the fellow's expertise in the utilisation of food consumption data, (2) to develop the fellow's knowledge on systematic review and meta-analysis methodologies and (3) to develop the fellow's capacity on how to weigh scientific evidence and to assess uncertainties of epidemiological research to be used under the risk assessment framework.

2.2. Activities/Methods

2.2.1. Identification of the subject matter and patterns of consumption of pulses in Greece and beyond

A single food is often known under several common names, whereas the use of similar names for different commodities is not uncommon. The foods belonging to the *Leguminosae* family (i.e. *fabaceae*, commonly known as the legume or pea family) are a good example of such inconsistencies, with terms such as 'pulses', 'legumes', 'dried vegetables' or 'grain legumes' being often used interchangeably by the scientific community. However, these terms characterise food items with different nutritional profiles, have distinct roles in eating habits of populations and have different implications for human and animal health, as well as for the environment. In the first module of this work programme, and to address *Problem formulation*, the fellow has explored, identified and characterised the definitions and inconsistencies among these terms through the consultation of selected sources (EFSA FoodEx2 and other food classification systems at the global or European level, food composition databases and dietary guidelines) that included food items belonging to the *Leguminosae* family in sufficient detail. In order to evaluate consumption patterns and trends for this food group, the fellow further consulted and worked with food intake data available for the global and European contexts. Food consumption data from Greece was used more extensively due to its traditional mediterranean eating pattern, of which pulses are an integral component, serving as a case study for a detailed insight into specific trends of consumption, types of pulses with higher

contributions to the overall intake and main characteristics of consumers. For these purposes, three levels of available data were considered: FAO Food Balance Sheets (FBS), compiling yearly country-level data on the available national food supply; Household Budget Surveys, providing information on food availability at the household level; and studies assessing food consumption at the individual level, in particular those following the EU Menu methodology, available at the EFSA Comprehensive European Food Consumption Database, further complemented by other data collections.

2.2.1.1. Definition of the exposure

The FoodEx2 – EFSA Food classification and description system for exposure assessment – is a comprehensive food classification system developed by EFSA (EFSA, 2015), including different grouping hierarchies, depending on the food domain of interest. The exposure-oriented hierarchy is designed to facilitate the grouping of food items for exposure calculations and the preferred hierarchy for reporting consumption data, being thus considered under the scope of this module. The EFSA Catalogue Browser was used as the main source to describe the specific terms and categorisation for the items under the *Leguminoase* family in FoodEx2. In this classification system, this food group is comprised within two major categories (hierarchy terms): 'Legumes, nuts, oilseeds and spices' and 'Vegetables and vegetable products'. The first is further divided into two hierarchy terms: 'Legumes', comprising two generic terms depending on whether the items are in the form of fresh ('Legumes fresh seeds') or dried seeds ('Pulses (dried legume seeds)'), with both categories including similar groups of several species of beans (including soyabeans), peas, lentils, lupins and other legumes (fresh seeds of horse gram, peanut, kersting's and bambara groundnut); and 'Processed legumes, nuts, oilseeds and spices', which is further divided into 'Canned or jarred legumes', including canned/jarred beans, peas, lentils and chickpeas, and 'Pulses flour', along with other core items pertaining to nuts, oilseeds and spices. On the other hand, the latter hierarchy term ('Vegetables and vegetable products') comprises the generic term of 'Legumes with pods', which is further divided into three core terms – 'Beans (with pods)', 'Lentils (with pods)' and 'Lentils (young pods)', with each term also including several species, along with other categories of vegetables and vegetable products. The major distinguishing aspect between these two main terms is that under the 'Legumes, nuts, oilseeds and spices', the fresh seeds of legumes without pods are considered; whereas in the category of 'Vegetables and vegetable products', only the fresh leguminous seeds within pods are included.

In order to further explore the definitions of this food group, other sources of food classification and description were consulted. A detailed comparison between the categorisation of FoodEx2 and other food classification systems for the species in the *Leguminosae* family is described in Appendix A. Briefly, FoodEx2, along with the FAO Definition and Classification of Commodities and FAO/WHO Codex classification are the most descriptive and comprehensive among the included classifications, describing a similar listing of species under the pulses category, with the exception of soyabeans, which are explicitly excluded from the second and are found under the category of oil-bearing crops. Additionally, it seems consensual among the selected sources that when the term pulses is used it refers exclusively to the dried legume seeds. In some cases, the category 'legumes' includes both the fresh and dried seeds and in other comprises specifically the fresh leguminous seeds, thus excluding pulses. The most evident discrepancies are related to the categorisation of the fresh legume seeds and processed legumes. Unlike in FoodEx2, other food classification systems do not distinguish legumes within pods from those in the shelled form. This might however be relevant from a chemical exposure perspective, since pods are fully exposed to pesticides, whereas the seeds are protected within the pod. In a similar way, processed legumes are described as a specific group in FoodEx2, an important distinction from a nutritional standpoint. However, it should be noted that different classification systems were developed to fill different needs and the inconsistencies identified might be a reflection of those rather than actual inconsistencies or disagreements.

2.2.1.2. Assessment of trends in the consumption of pulses worldwide, in Europe and Greece

According to the OECD-FAO Agricultural Outlook 2020–2029 (OECD/FAO, 2020), about 22 g/day/per capita of pulses were available for consumption in 2020, globally. The values were highest in Africa and Latin America (about 30 g/day/per capita), followed by Asia (19 g/day/per capita) and North America (16 g/day/per capita) and lowest in Europe (8 g/day/per capita) and Oceania (5 g/day/per capita). The overall supply of pulses had a slow but steady decline in both high- and low-income regions since the early 1960s, reaching a plateau between the early 1990s and 2000s. No major changes are foreseen in the global per capita availability of pulses for the next decade (OECD/FAO,

2020). However, pulses are expected to regain importance in the diets and farming systems of Europe as a whole and in the European Union in particular (OECD/FAO, 2020). After consulting the available data from FAO FBS (FAOSTAT, 2021), in 2010, the per capita supply of pulses in Europe was estimated at 6.8 g/day, with a slow but steady increase since then. This value is expected to increase to 9.5 g/day per capita in 2029. Between 1961 and 2018, peas were the most available pulse in Europe, with an average of 2.5 g/day per capita in 2018, followed by beans with 1.7 g/day per capita in the same year. However, the supply available for consumption of both of these types of pulses has been slowly decreasing throughout time, while for other pulses (including broad beans, horse beans, chickpeas, cowpeas, pigeon peas, lentils, bambara beans, vetches, lupins, flour and bran of pulses), it has been steadily increasing. Due to the nature of categorisation of data in FAO FBS, it was not possible to understand if this is due to higher levels in the supply of a single or more types of pulses.

Focusing on a National context, data available at the FAO FBS for Greece, the Greek Household budget surveys and original studies on food consumption at the individual level were consulted and used. Table 1 shows the trends in the average per capita supply (g/day) in Greece by decade and the respective average share in the total caloric and protein supply for that time period, according to FAO FBS. Similarly to the observed for the global context, the per capita supply of pulses in the country has been steadily declining since the early 1960s from 21.5 g/day to 13.4 g/day in the 2010s. In the 1960s, pulses represented 5.1% of the total protein supply in Greece and 2.5% of the total supply of calories intake (Table 1). Six decades later, in the 2010s, these values are about half of those previously reported. These trends were also in line with those described using data from the Greek Household budget surveys.

Table 1: Average per capita per day supply of pulses and its contribution to the supply of protein and calories in Greece (1961–2018)

	1961–1970	1971–1980	1981–1990	1991–2000	2001–2010	2011–2018
Average consumption (g)	21.5	18.6	14.4	13.9	13.0	13.4
% of total protein supply	5.1	3.8	2.8	2.6	2.5	2.7
% of total energy supply	2.5	1.9	1.4	1.3	1.2	1.3

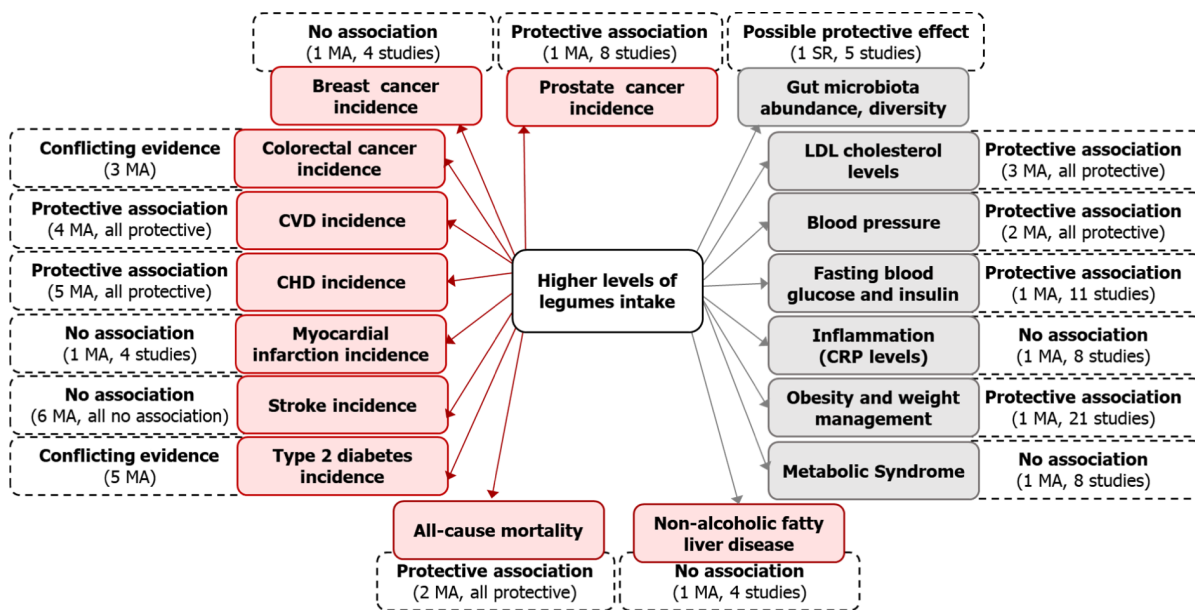
All estimates were calculated using data available at the FAO Food Balance Sheets for Greece (FAOSTAT, 2021).

At the individual level, Greek food consumption data available at the EFSA Comprehensive Food consumption database were retrieved. The number of individuals reporting any level of intake of pulses was relatively low in all age groups. For the pulses' food group as a whole, the percentage of consumers was 18.2% in adolescents, 15.8% in adults and 23.7% in the elderly in 2015. The average consumption was higher among the elderly [mean (sd): 6.73 (15.08) g/day] and lower among adolescents [mean (sd): 4.72 (14.13) g/day]. Lentils were the most consumed pulse among adolescents (9.1%), while beans were reported by a higher number of adult and elderly consumers (10.4% and 11.3%, respectively). Taking into account the Greek Dietary Guidelines (Kastorini et al., 2019), which recommend an amount of at least 450 g of pulses per week for adolescents and adults, the results provided in this study show that regardless of the age group, the Greek population is currently characterised by low levels of intake of pulses, with an average of about 33 g/week among adolescents, 34 g/week among adults and 47 g/week among the elderly.

2.2.2. Understanding of the current state of the art: health outcomes associated with the consumption of legumes

The second module of the work programme consisted in conducting of a comprehensive literature search to describe the health outcomes associated with the consumption of pulses (hazard identification) and to identify existing dose–response meta-analyses (hazard characterisation). For this purpose, the fellow searched PubMed and other literature sources in order to identify systematic reviews and meta-analyses addressing the association between the dietary intake of pulses and health outcomes, considering either the whole food group or specific types. Despite the main interest of the present work programme being legumes in their dried form, i.e. pulses, after a preliminary search it was evident that in the majority of studies there was no distinction between subtypes, with the food group of legumes being assessed as an aggregated term, including fresh and dried seeds. For this reason, the term legumes was used in the sections of this report aiming to address the epidemiological evidence on the association between this food group and health outcomes (Sections 2.2.2 and 2.2.3).

In addition, studies focusing exclusively on the health effects of the corresponding nutrients and other compounds (e.g. proteins, fibre, isoflavones) were excluded, in an attempt to capture interactions among food constituents and approach holistically the association between food intake and disease risk. Another aim of this module was to gather the necessary evidence to allow the selection of a health outcome to be studied in the context of a dose–response meta-analysis in module 3. For this reason, studies assessing exclusively the health effects of soybeans and soy products were also excluded, since these have already been extensively reviewed elsewhere (46 systematic reviews and meta-analyses identified in a preliminary search). For the association between the intake of legumes and pulses (excluding soy) and health outcomes, including hard clinical outcomes and potential intermediate factors (surrogate endpoints), 44 systematic reviews and/or meta-analyses were identified. An overview of the conclusions resulting from these studies is described in Figure 1.



The associations with hard clinical endpoints are highlighted in red and with intermediate factors in gray. For each outcome, the direction of the association with the intake of legumes is indicated. For those cases where only one meta-analysis was performed, the number of included studies is presented in parenthesis, whereas when two or more meta-analyses were available for the same association, the agreement or disagreement in their conclusions is noted also in parenthesis. MA: meta-analysis; SR: systematic review; CVD: cardiovascular disease; CHD: coronary heart disease; CRP: C-Reactive Protein.

Figure 1: Consumption of legumes and related health outcomes, according to the results from the identified systematic reviews and meta-analyses

Through the literature search performed, 11 publications with a dose–response meta-analysis on the association between the intake of legumes and one or more hard clinical endpoints were identified. Considering the most recent meta-analyses conducted for each outcome, a statistically significant dose–response association was reported for Type 2 diabetes and CVD only, albeit with a limited number of included studies. The strengths and limitations of these studies were assessed and discussed with the supervisor and other team members, aiming to increase the fellow’s capacity to critically appraise the available evidence and to inform the next steps of the work programme.

In order to enhance the relevance and interpretation of findings in terms of health impact, the fellow has consulted the Global Burden of Disease Study (GBD, 2019) to identify estimates of disease burden associated with each hard clinical endpoint identified in Figure 1. Disability adjusted life years (DALYs), combining years of potential life lost due to premature mortality and years lived with a disability of specified severity and duration, reported in the Global Burden of Disease Study 2019 (GBD, 2019) were used as the estimate of interest. In 2019, CVD was responsible for 393 million (95% UI 368–417) DALYs and was the overall leading cause of disease burden – specifically for CVD subtypes, coronary heart disease reached a total 182 million (95% UI 170–194) DALYs and stroke a total of 143 million (95% UI 133–153). Type 2 diabetes was second with 66.3 million (95% UI 55.5–79.0) global DALYs in 2019, followed by colorectal cancer with 24.3 million (22.6–25.7) DALYs.

The final step of this module consisted in selecting the health outcome to be studied in the context of a dose–response meta-analysis. For this purpose, five criteria were defined and applied: (1) hard clinical endpoints with a plausible association with the intake of legumes; (2) outcomes with no previous or recent dose–response meta-analysis exploring their association with the intake of legumes; (3) outcomes with available estimates of disease burden (i.e. DALYs); (4) outcomes with an established positive association with the intake of red meat and red meat products, considering the substitution of red meat with plant-based protein sources as the wider scope of this work programme; (5) availability of several publications for the association between the health outcome and the intake of legumes, so as to increase the precision of the effect measures. For all the identified hard clinical endpoints with a plausible association with the consumption of legumes (Figure 1) (criterion 1), a relatively recent dose–response meta-analysis was performed (criterion 2) and with the exception of all-cause mortality, all the identified outcomes had an associated estimate of disease burden (criterion 3). Colorectal cancer and CVD have an established positive association with the intake of red meat and a recent dose-response meta-analysis has also linked the intake of red meat to the development of type 2 diabetes (Yang et al., 2020) (criterion 4). However, two dose-response meta-analyses on this association were published in 2020 and 2021, limiting the need to conduct a new synthesis of findings. CVD and colorectal cancer were therefore regarded as the most suitable outcomes for the investigation proposed within the context of the present work plan. Considering the availability of publications for the association between these health outcomes and the intake of legumes (criterion 5), preliminary literature searches were conducted. For colorectal cancer, the preliminary search retrieved 141 papers published after the most recent update (Schwingshackl et al., 2018), all with a cross-sectional or retrospective design. On the other hand, a preliminary search on the association between the intake of legumes and CVD after the most recent update (Viguiliouk et al., 2019) retrieved 557 publications, with original research addressing the association of interest including prospective and retrospective study designs. Prospective studies are generally considered to rank higher than retrospective approaches in the hierarchy of evidence, thus improving the strength and robustness of pooled effect measures. For this reason, overall CVD and its different subtypes were selected as the health outcomes to be explored within the context of a dose–response meta-analysis.

2.2.3. Assessment of dose–response associations between the consumption of legumes and the selected health outcome

The third module of the work programme consisted in the evaluation of dose–response associations between the intake of legumes and the health outcomes selected in module 2 – CVD and subtypes – in the context of a systematic review and meta-analysis. In an initial stage, the fellow consulted the guidelines from the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2019) and the EFSA guidance document on the application of systematic review methodology to food and feed safety assessments (EFSA, 2010), along with other relevant literature, to support the methodological design and procedures of the study. After a thorough discussion with the supervisor and the supporting team on methods and major aspects to be considered, the study protocol was prepared following the PRISMA-P guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocols) (Moher et al., 2015) and registered in PROSPERO (International prospective register of systematic reviews), an open access online database of systematic review protocols in health and social care, so as to ensure transparency and to enable *a posteriori* comparisons of reported methods with those initially planned (PROSPERO registration ID: CRD42021247565).

2.2.3.1. Literature search and selection of studies

In an initial step, several literature databases were considered for literature search. After assessing their scope and characteristics, taking into account comprehensiveness and complementarity as criteria, PubMed, Scopus and Web of Science were considered and searched since inception in order to retrieve potential eligible studies. For this purpose, several search expressions were developed and tested, in accordance to each database format and specific controlled vocabulary thesaurus. Terms designating the species belonging to the *Leguminosae* family were included, using both their common and scientific names. The final search expression was designed so as to ensure an adequate balance between specificity and sensitivity, however favouring a higher sensitivity in order to capture the highest number of eligible studies. After performing several sensitivity analyses, broader terms such as 'dietary protein', 'plant protein' and 'Mediterranean diet' were included, resulting in a total of 14 374 publications for reviewing, after the removal of duplicates. Study selection was then developed in three

stages: (1) title/abstract review, (2) full-text review and (3) evaluation of data in the format required for synthesis, applying the selection criteria defined at the protocol stage, and performed independently by the fellow and other member of the scientific supporting team. Lastly, a back and forward citation tracking of eligible studies was conducted, in order to ensure the completeness of study selection.

2.2.3.2. Data extraction and dose calculation

A data extraction template was created to ensure a standardised approach to data extraction, including all the relevant variables to allow a robust analysis. The data were extracted independently by the fellow and another member of the scientific supporting team for validation purposes. For each exposure category, the median or mean intake level was extracted, depending on the availability of the data provided by the authors. In cases where this information was reported in intervals of consumption, the midpoint of each exposure strata was used and when the highest and lowest exposure category intervals were open-ended, a value that was respectively 20% higher and lower than the closest cut point was used as boundary. In cases where the intake levels were reported only as number of servings, a standard portion size was assigned. Finally, all the dose levels were converted into a standard unit, in order to allow comparability between estimates for the dose-response assessment.

2.2.3.3. Risk of bias assessment

For the assessment of the risk of bias in observational research, the use of two possible frameworks was initially considered – the Risk of Bias in Non-randomised Studies of Interventions (ROBINS-I) and the Risk of Bias in Non-randomised Studies of Exposures (ROBINS-E) – and discussed, comparing advantages and limitations. Both are based on the Cochrane risk of bias (RoB) tool for randomised trials and follow the premise that an observational study should be compared to a hypothetical randomised trial to identify potential biases, comprising similar domains and signalling questions. However, ROBINS-E, even if apparently more suited for the nature of the included research in the present study, is still currently under development and lacks an accompanying guidance document. On the contrary, ROBINS-I is a well-established tool with a detailed and comprehensive guidance document published (Sterne et al., 2016). For this reason, ROBINS-I was selected for the RoB assessment. In accordance with the guidance document, the tool was tailored to address the specific research question accommodating the nature and methodology of studies selected to be considered in the analysis. Table 2 summarises the assumptions considered in the RoB assessment across the seven domains covered in ROBINS-I. Each domain was classified as having a low, moderate, serious or critical risk of bias, with a derived overall assessment for each study. The appraisal was performed in parallel by the fellow and another researcher using a standardised template specifically created for this purpose, in order to ensure the repeatability of the process and ultimately the comparability of the assessments.

Table 2: Segments of the protocol adopted for the RoB assessment

ROBINS-I domains	Criteria
Bias due to confounding	Factors mandatory to be considered in the analysis so as to judge a study at low RoB were sex, age, energy intake and BMI; mandatory factors for a moderate RoB judgement were age, sex, BMI and measures of dietary factors/patterns (in the absence of energy intake) or age, sex, energy intake, physical activity and clinical CVD risk factors (in the absence of BMI).
Bias in selection of participants	Selection of eligible participants must not be related to either legumes intake or to CVD.
Bias in classification of interventions	Studies were judged to be at a low RoB for exposure misclassification if using a validated dietary assessment method and quantifying the dose of exposure (i.e. portion consumed). If one of these conditions was not satisfied, the study would be judged to be at moderate RoB.
Bias due to deviations from intended interventions	There should be no concern about departure from intended exposure due to the long-term stability of dietary patterns. A threshold of > 10 years of mean follow-up was considered at moderate RoB for possible change in exposure since the beginning of the study, if suggested that factors conditioning the change were not at random.

ROBINS-I domains	Criteria
Bias due to missing data	Studies with less than 10% of missing data in the exposure, outcome or covariates were considered at low risk of bias. If more than 10% of the data was missing for one of these domains, the studies were judged to be at moderate RoB.
Bias in outcome measurement	Studies were judged to be at low risk of bias if using valid methods for the outcome assessment (e.g. medical records, self-reported diagnosis of CVD with external validation, death registries/autopsies). Studies only partially validating the outcome assessment were judged to be at moderate RoB.
Bias in selection of the reported results	Suggestion that published results have not been selected and clear reporting of statistical methods. If the analysis described in the methods section does not correspond entirely to what is presented in the results section, but unlikely to mean selective reporting on the basis of the results, the study was judged to be at moderate RoB.
Overall RoB judgement	If all domains were at low risk of bias, the overall risk was considered low; if at least one domain was found at moderate risk of bias, the overall risk was considered moderate. The same approach was applied for serious and critical risk of bias.

2.2.3.4. Dose–response analysis

Data were analysed following two approaches: firstly, a traditional meta-analysis was conducted comparing highest to lowest intakes. A random effects model was applied and the heterogeneity across included studies was assessed using the I^2 statistic. In a second step, the dose–response analysis was performed. The methodology developed by Orsini et al. (2012) and Crippa et al. (2019) was used to explore the shape of the relationship between the intake of legumes and CVD (overall and separately for CHD and stroke). The advantage of the dose–response model applied (one-stage approach) is that studies with as few as two levels of exposure can be considered through the application of a weighted mixed-effects model. The one-stage approach was carried out, using a restricted cubic spline model with three knots at fixed percentiles (10, 50 and 90%), under no *a priori* assumption on the shape of the curve. The estimates were pooled using the restricted maximum likelihood method in mixed-effects models. The possible presence of publication bias was verified through visual inspection of funnel plots. Several sensitivity and subgroup analyses were conducted to understand if the associations varied with different study characteristics. Finally, a manuscript was prepared following the PRISMA guidelines to be submitted for publication.

2.2.4. Additional activities

2.2.4.1. Preliminary literature review on anti-nutrients present in pulses

There is a general consensus that a diet rich in plant-based foods has an important role in the prevention and reduction of disease. However, despite being rich sources of micro- and macronutrients, plant-based foods contain significant concentrations of other compounds for which the health effects are not yet fully understood. These compounds are often referred to as antinutrients (also bioactive compounds or phytochemicals), as they are thought to restrict bioavailability of key nutrients. The fellow has completed a preliminary narrative review on the presence of antinutrients in pulses, aiming to identify potential differences between species and possible impact in human nutrition, as well as to acknowledge current research gaps. Briefly, pulses are generally a rich source of antinutrients, including lectins, enzyme inhibitors, phytates, oxalates, phytoestrogens, saponins, alkaloids and oligosaccharides, and their content in these factors may widely vary between different subtypes, as well as between different species within the same subtype (Alcázar-Valle et al., 2020; Mayer Labba et al., 2021). The evidence on the potential health effects of antinutrients in pulses is mixed, with some studies showing they may condition the adequate absorption of nutrients and have other negative health effects (Fredlund et al., 2006; Petroski and Minich, 2020), while others demonstrate their potential in disease prevention, in particular for phytoestrogens, lectins and phytates (Gautam et al., 2020; Yamagata and Yamori, 2021; Silva and Bracarense, 2016). However, the concentration of these compounds in pulses is significantly reduced through traditional processing such as soaking, sprouting, fermenting, boiling and autoclaving (Luo and Xie, 2013; Bento et al., 2021), and there is yet not sufficient evidence showing their potential effects on human health (Petroski and Minich, 2020). The interaction between antinutrients and other constituents in food matrixes is also not well established (Petroski and Minich, 2020).

Other additional activities completed during the fellowship period are reported in Appendix B.

3. Conclusions

The work programme at the Dept. of Hygiene, Epidemiology and Medical Statistics of the School of Medicine, National and Kapodistrian University of Athens provided the fellow the opportunity to develop important skills within critical aspects of the risk assessment framework. Using the *Leguminosae* family as a case study, the fellow has become familiar with a different range of food classification systems, food composition tables and dietary guidelines and has significantly improved her knowledge in the utilisation of these tools for identifying inconsistencies across the different sources and establishing a definition for the exposure of interest (an important step in **hazard identification**). The fellow has extensively explored the EFSA Comprehensive European Food Consumption Database and has improved her knowledge on the FAO Food Balance Sheets and the Household Budget Surveys databases in order to assess temporal trends in the consumption of pulses worldwide, in Europe and Greece. This has allowed the fellow to substantially advance her skills in the utilisation of food consumption databases for the assessment and monitoring of habitual dietary intakes at the individual, household and country level. The differences and complementary aspects across these dietary assessment methods were also analysed taking into account sources of potential bias, in order to improve the fellow's knowledge and skills on the critical appraisal of available food consumption data (a critical aspect of **exposure assessment**). In order to advance the skills of the fellow on **hazard characterisation**, the fellow has conducted a literature search to identify and summarise the evidence that investigated the association between different health outcomes and the consumption of legumes. This has enhanced the fellow's knowledge on how to identify and assess epidemiological evidence addressing the complex relationship between diet and disease. The fellow has also increased her expertise on systematic review and meta-analysis methodologies, by preparing a systematic review and a dose-response meta-analysis in accordance with the guidelines from Cochrane and EFSA. The fellow has further acquired knowledge in the appraisal of risk of bias in observational research and improved expertise on weighing of scientific evidence in nutritional epidemiology studies related to hazard characterisation. The EU-FORA fellowship programme and the stay at the University of Athens was an extremely rewarding experience for the fellow, both in terms of professional and personal development.

References

- Alcázar-Valle M, Lugo-Cervantes E, Mojica L, Morales-Hernández N, Reyes-Ramírez H, Enríquez-Vara JN and García-Morales S, 2020. Bioactive compounds, antioxidant activity, and antinutritional content of legumes: a comparison between four phaseolus species. *Molecules*, 1, 25, 3528.
- Bento JAC, Ribeiro PRV, Alexandre e Silva LM, Alves Filho EG, Bassinello PZ, de Brito ES, Caliani M and Soares Júnior MS, 2021. Chemical profile of colorful bean (*Phaseolus vulgaris* L) flours: changes influenced by the cooking method. *Food Chemistry*, 15, 129718.
- Crippa A, Discacciati A, Bottai M, Spiegelman D and Orsini N, 2019. One-stage dose-response meta-analysis for aggregated data. *Statistical Methods in Medical Research*, 28, 1579–1596.
- EFSA (European Food Safety Authority), 2010. Application of systematic review methodology to food and feed safety assessments to support decision making. *EFSA Journal* 2010;8(6):1637, 90 pp. <https://doi.org/10.2903/j.efsa.2010.1637>
- EFSA (European Food Safety Authority), 2015. The food classification and description system FoodEx2 (revision 2). *EFSA Supporting Publication* 2015;12(5):EN-804, 90 pp. <https://doi.org/10.2903/j.efsa.2015.en-804>
- FAOSTAT, 2021. Food balance sheets. Available online: <https://www.fao.org/faostat/en/#data/FBS>
- Fredlund K, Isaksson M, Rossander-Hulthén L, Almgren A and Sandberg A-S, 2006. Absorption of zinc and retention of calcium: dose-dependent inhibition by phytate. *Journal of Trace Elements in Medicine and Biology*, 20, 49–57.
- Gautam AK, Sharma D, Sharma J and Saini KC, 2020. Legume lectins: potential use as a diagnostics and therapeutics against the cancer. *International Journal of Biological Macromolecules*, 1, 474–483.
- GBD (Global burden of disease), 2019. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*, 396, 1204–1222.
- Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ and Welch VA, 2019. *Cochrane Handbook for Systematic Reviews of Interventions*, 2nd Edition. John Wiley & Sons, Chichester (UK).
- Kastorini C-M, Critselis E, Zota D, Coritsidis AL, Nagarajan MK, Papadimitriou E, Belogianni K, Benetou V and Linos A, 2019. National Dietary Guidelines of Greece for children and adolescents: a tool for promoting healthy eating habits. *Public Health Nutrition*, 22, 2688–2699.

- Luo YW and Xie WH, 2013. Effect of different processing methods on certain antinutritional factors and protein digestibility in green and white faba bean (*Vicia faba* L.). *CyTA-J Food*, 11, 43–49.
- Mayer Labba I-C, Frøkiær H and Sandberg A-S, 2021. Nutritional and antinutritional composition of fava bean (*Vicia faba* L., var. *minor*) cultivars. *Food Research International*, 140.
- Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P and Stewart LA, 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4, 1.
- OECD/FAO, 2020. OECD-FAO Agricultural Outlook 2020–2029. Rome/OECD Publishing, Paris, FAO.
- Orsini N, Li R, Wolk A, Khudyakov P and Spiegelman D, 2012. Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *American Journal of Epidemiology*, 175, 66–73.
- Petroski W and Minich DM, 2020. Is there such a thing as “anti-nutrients”? A narrative review of perceived problematic plant compounds. *Nutrients*, 12, 2929.
- Schwingshackl L, Schwedhelm C, Hoffmann G, Knüppel S, Laure Preterre A, Iqbal K, Bechthold A, De Henauw S, Michels N, Devleeschauwer B, Boeing H and Schlesinger S, 2018. Food groups and risk of colorectal cancer. *International Journal of Cancer*, 142, 1748–1758.
- Silva EO and Bracarense AP, 2016. Phytic acid: from antinutritional to multiple protection factor of organic systems. *Journal of Food Science*, 81, R1357–R1362.
- Sterne JAC, Higgins JPT, Elbers RG, Reeves BC, 2016. Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I): detailed guidance, updated 12 October 2016.
- Viguiouk E, Glenn AJ, Nishi SK, Chiavaroli L, Seider M, Khan T, Bonaccio M, Iacoviello L, Mejia SB, Jenkins DJA, Kendall CWC, Kahleová H, Rahelić D, Salas-Salvadó J and Sievenpiper JL, 2019. Associations between dietary pulses alone or with other legumes and cardiometabolic disease outcomes: an umbrella review and updated systematic review and meta-analysis of prospective cohort studies. *Advances in Nutrition*, 10(Suppl. 4), S308–S319.
- Yamagata K and Yamori Y, 2021. Potential effects of soy isoflavones on the prevention of metabolic syndrome. *Molecules*, 26, 5863.
- Yang X, Li Y, Wang C, Mao Z, Zhou W, Zhang L, Fan M, Cui S and Li L, 2020. Meat and fish intake and type 2 diabetes: dose-response meta-analysis of prospective cohort studies. *Diabetes Metabolism*, 46, 345–352.

Abbreviations

CVD	Cardiovascular disease
DAFNE	Data Food Networking
DALYs	Disability-adjusted Life Years
EU-FORA	European Food Risk Assessment Fellowship Programme
FAO	Food and Agriculture Organization
FBS	Food balance sheets
OECD	Organisation for Economic Co-operation and Development
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	International prospective register of systematic reviews
RoB	Risk of bias
ROBINS-I	Risk of Bias in Non-randomised Studies of Interventions

Appendix A – Comparison between EFSA FoodEx21 and other food classification systems for the food group of species in the Leguminosae family

Food classification system	FoodEx2 hierarchy term	FoodEx2 generic term	Terms and description used in the respective food classification system for equivalent categories of the FoodEx2* classification system
FAO Definition and Classification of Commodities²	Legumes	Pulses (dried legume seeds)	<u>Pulses and derived products</u> : annual leguminous crops yielding from one to 12 grains or seeds of variable size, shape and colour within a pod. The term 'pulses' is limited to crops harvested solely for dry grain. Also includes processed pulses like flour and bran of pulses and explicitly excludes soyabeans.
	Vegetables and vegetable products	Legumes with pods	<u>Vegetables and Derived Products</u> : pulses belong to this group when harvested green; the list includes green beans, green peas (edible-podded peas or sugar peas), green broad beans and string beans, mostly for shelling.
FAO/WHO Codex Classification of Foods and Feeds³	Legumes	Pulses (dried legume seeds)	<u>Pulses</u> : derived from the mature seeds, naturally or artificially dried of leguminous plants, consumed after processing or household cooking. Included under the broader category of vegetables. Includes soyabeans.
	Vegetables and vegetable products	Legumes with pods	<u>Legume vegetables</u> : derived from the succulent seed and immature pods of leguminous plants; may be consumed as whole pods or as the shelled product. Included under the broader category of vegetables. Includes soyabeans.
WCO Harmonized Commodity Description and Coding System⁴	Legumes	Pulses (dried legume seeds)	<u>Dried leguminous vegetables</u> : includes only shelled, whether or not skinned or split, dried seeds of leguminous plants. Covers the same subgroups and roughly the same species, with some exceptions (e.g. lupins and vetches not mentioned). Soyabeans are not mentioned.
	Vegetables and vegetable products	Legumes with pods	<u>Leguminous vegetables</u> : includes species shelled or unshelled, in their fresh or chilled form. Only specifies generic species of beans and peas. Soyabeans are not mentioned.
DAFNE Food Classification System⁵⁺	Legumes	Pulses (dried legume seeds)	<u>Pulses</u> : the nomenclature varies with country, but most use the term 'dried' peas/beans/pulses or 'dried vegetables'. Presents limited detail on included specific species. Some countries include canned pulses in this category (France and Italy).
	Processed legumes, nuts, oilseeds and spices	Canned/ jarred legumes	<u>Processed vegetables</u> : including canned, frozen, boiled/cooked food items from the leguminosae family. The reporting of this category varies widely according to the country. Presents limited detail on included specific species.
	Vegetables and vegetable products	Legumes with pods	<u>Fresh vegetables</u> : generally, only the fresh forms of the leguminous plants seeds or pods are included. Presents limited detail on included specific species.

1: EFSA (European Food Safety Authority), 2015. The food classification and description system FoodEx2 (revision 2). EFSA supporting publication 2015;EN-804, 90 pp.

2: FAO (Food and Agriculture Organization of the United Nations), 1994. Definition and classification of commodities, 4. Pulses and derived products.

3: Codex Alimentarius Commission, 1993. Codex Classification of Foods and Animal Feeds. Codex Alimentarius, 2.

4: WCO (World Customs Organization), 2017. The Harmonized Commodity Description and Coding Systems (HS).

5: European Commission, DG SANCO, 2005. The DAFNE food classification system. Operationalisation in 16 European countries. Services of the European Commission, Luxembourg.

*: FoodEx2 categories without correspondence in the designated food classification system are omitted.

†: Classification system of data collected from the National Household Budget Surveys in Europe.

Appendix B – Additional activities

Next to the training provided in the context of the EU-FORA fellowship programme, including the five EU-FORA modules organised by EFSA (Italy), AGES (Austria), BfR (Germany) and EFET (Greece), the fellow also had the opportunity to attend and participate in other training activities covering a wide range of topics, as described below.

	Description of activity/training	Tutor	Date
1	Observer and commentator on presentations of the students' on dietary survey data available in Greece, in the context of the Nutritional Epidemiology course of the MSc in Epidemiology – Research methodology in biomedical sciences, clinical practice and public health of the School of Medicine, National and Kapodistrian University of Athens (NKUA)	–	March 2021
2	Attendance to the WHO/Europe NCD Office Seminar for Early Career Researchers – 'Global health career development'	Dr Gauden Galea	26 March 2021
3	Lecture 'Analysing nutrient and food data' to the students of the MSc in Epidemiology – Research methodology in biomedical sciences, clinical practice and public health of the School of Medicine of the National and Kapodistrian University of Athens	–	1 April 2021
4	Attendance to the seminar 'Tackling Non-communicable diseases and promoting Public Health in WHO/Europe', organised by the DeHMS/NKUA	Dr João Breda	12 April 2021
5	Attendance to the WHO/Europe NCD Office Seminar for Early Career Researchers – 'Novel opportunities of interdisciplinary approaches'	Dr Afton Halloran	21 April 2021
6	Attendance to the WHO/Europe NCD Office Seminar for Early Career Researchers – 'Alcohol and Cancer: Evidence for Action'	Prof. Linda Bauld	25 May 2021
7	Attendance to the seminar 'From epidemiologic evidence to formulation and implementation of nutrition policy: the case of trans-fatty acids', within the scope of the Nutritional Epidemiology Course organised by the DeHMS/NKUA	Dr Theodora Mouratidou/ Dr Sandra Caldeira	27 May 2021
8	Attendance to the WHO/Europe NCD Office Seminar for Early Career Researchers – 'Corporate political activity as a commercial determinant of health: a case study and methodological considerations'	Dr Kathrin Lauber	28 June 2021
9	Attendance to the Conference 'Future steps to tackle obesity – digital innovations into policy and actions', organised by the Portuguese Ministry of Health in the context of the Portugal's Presidency of the Council of the EU	–	29 June 2021
10	Attendance to the 'International workshop on risk assessment of combined exposure to multiple chemicals', promoted by EFSA	–	18-20 October 2021
11	Attendance to the Conference of the WHO Collaborating Centre for Nutrition and Childhood Obesity by the National Institute of Health Dr Ricardo Jorge, Portugal	–	19 October 2021
12	Attendance to the 9th Virtual Panhellenic Conference of Greek Lipid Forum 'Current trends in the field of Lipids'	–	22 October 2021
13	Completion of the Cochrane Interactive Learning Course: 'Conducting an intervention review', promoted by Cochrane through the People Services Department	–	November 2021
14	Attendance to the WHO seminar 'Improving digital food environments: why do we need to act now?'	–	7 December 2021