

A Standardised Approach Towards PROving the Efficacy of Foods and Food Constituents for Health CLAIMs (PROCLAIM): Providing Guidance

Alison M. Gallagher¹, Gert W. Meijer², David P. Richardson³, Virginie Rondeau⁴, Maria Skarp⁵, Marianne Stasse-Wolthuis⁶, Guy C. Tweedie⁷ and Renger Witkamp^{8,9}

1. Northern Ireland Centre for Food and Health, School of Biomedical Sciences, University of Ulster, Coleraine BT52 1SA, UK
2. Unilever R&D, Olivier van Noortlaan 120, 3133 AT Vlaardingen, The Netherlands
3. DPR Nutrition Limited, 34 Grimwade Avenue, Croydon, Surrey CR0 5DG, UK
4. Division of Biostatistics, Unité 897, Institut National de la Santé et de la Recherche Médicale, University Victor Segalen of Bordeaux 2, Bordeaux F-33076, France
5. ILSI Europe, Avenue E. Mounier 83, Box 6, 1200 Brussels, Belgium
6. Stasse Consultancy, Udenhout, The Netherlands
7. Danone Spain, Calle Corcega 590, Barcelona, Spain
8. Division of Human Nutrition, Wageningen University, Bomenweg 2, 6703 HD, Wageningen, The Netherlands
9. TNO Quality of Life, Utrechtseweg 48, Zeist, The Netherlands



Commissioned by the
ILSI Europe Functional Foods Task Force

Correspondence: ILSI Europe a.i.s.b.l. - Avenue E. Mounier 83, Box 6 - 1200 Brussels - Belgium
email publications@ilsieurope.be - Fax: +32 2 762 00 44

Table of Contents

General approaches to claim substantiation	S18
Health claim definitions and regulatory frameworks	S18
Criteria for the scientific substantiation of health claims	S19
Building a scientific concept	S20
Establishing benefit to health	S20
Representative study group: the concept of health as a continuum	S20
Selecting appropriate study types and designs	S21
Choice of risk factors/markers	S22
Biological plausibility	S24
Weighing the totality of the evidence	S24
Conclusions and key recommendations	S26
Acknowledgements	S26

(Received 1 December 2010 – Revised 12 April 2011 – Accepted 15 May 2011)

Key words: Proving efficacy: Substantiation: Foods or food constituents: Health claims: International Life Sciences Institute Europe

Correspondence: ILSI Europe a.i.s.b.l., Avenue E. Mounier 83, Box 6, 1200 Brussels, Belgium, fax +32 2 762 00 44, email: publications@ilsieurope.be

Abbreviations: EFSA, European Food Safety Authority; FDA, Food and Drug Administration; ILSI, International Life Sciences Institute; PASSCLAIM, Process for the Assessment of Scientific Support for CLAIMs on foods; RCT, randomised controlled trials

Diet is well known to have beneficial health properties that extend beyond traditionally accepted nutritional effects. The approach involved in elucidating these beneficial physiological effects is becoming more important, as reflected by increasing research being undertaken. With growing consumer awareness of foods and food constituents and their relationship to health, the key questions for regulators, scientists and the food industry continue to relate to: (1) how consumers could be protected and have confidence that the health claims on foods are well supported by the evidence; (2) how research on physiological effects of food (constituents) and their health benefits could be stimulated and supported; (3) how research findings could be used in the development of innovative new food products. The objectives of this paper are to provide a set of recommendations on the substantiation of health claims for foods, to develop further guidance on the choice of validated markers (or marker patterns) and what effects are considered to be beneficial to the health of the general public (or specific target groups). Finally, the case for developing a standardised approach for assessing the totality of the available scientific data and weighing the evidence is proposed.

The importance of developing a standardised evidence-based approach for establishing food–health relationships has been addressed previously by the ‘Process for the Assessment of Scientific Support for CLAIMs on foods’ (PASSCLAIM)⁽¹⁾. Following exchanges across industry, regulators and academia⁽²⁾, the objective of this paper was to provide a set of recommendations on the scientific substantiation of health claims for foods, to develop further guidance on the choice of validated (generally accepted) markers (or marker patterns) and on what effects are considered to be beneficial to the health of the general public (or specific target groups). (The current paper focuses on food and food constituents with health benefits that are aimed at the generally healthy population. Clinical nutrition meets the particular nutritional requirements of individuals affected by, or who are malnourished due to, a specific disease or condition. Communication on these specific benefits is not regulated under the European Nutrition and Health Claim Regulation. For this reason, clinical nutrition is outside the scope of this publication.)

General approaches to claim substantiation

Health claim definitions and regulatory frameworks

Although definitions of health claims are slightly different around the world (e.g. European Parliament and Council⁽³⁾; US Food and Drug Administration (FDA)⁽⁴⁾; Codex Alimentarius⁽⁵⁾), usually a distinction is made between ‘nutrient function’ claims, ‘other function’ or ‘enhanced function’ claims and ‘reduction of disease risk’ claims. For instance, in the European Union, a ‘health claim’ is any claim that states, suggests or implies that a relationship exists between a food category, a food or one of its constituents and health. A ‘reduction of disease risk’ claim is defined as any health claim that states, suggests or implies that the consumption of a food category, a food or one of its constituents significantly reduces a risk factor in the development of a human disease. Claims that do not refer to disease risk reduction relate to a positive contribution to health, to the improvement of a function or to preserving health. Codex alimentarius defines that a ‘nutrient function’ claim describes the role of a nutrient intended to affect normal body structure or function⁽⁵⁾. An ‘other function’ claim (or ‘enhanced function’ claim) is defined as a claim concerning a specific beneficial effect of the consumption of foods or their constituents (not including

nutrients), in the context of the total diet, on a physiological function or biological activity⁽⁵⁾.

The aim of regulatory frameworks in different regions of the world is to ensure confidence in health claims on foods by requiring that all authorised claims are scientifically substantiated, and thereby to promote innovation and to achieve a high degree of consumer protection.

In Europe, regulation 1924/2006 applies to nutrition and health claims made in commercial communications and sets out the conditions for their use, establishes a system of scientific evaluation and creates European Community lists of authorised claims⁽³⁾. It requires that the claims are based on ‘generally accepted scientific evidence’ and that they are well understood by the ‘average consumer’.

The European Food Safety Authority (EFSA) Panel on Dietetic Products, Nutrition and Allergies assesses health claims under Article 13.1 (‘general function’ health claims), Article 13.5 (‘new function’ health claims, based on newly developed scientific evidence and/or that include a request for protection of proprietary data) or Article 14 (‘reduction of disease risk’ claims and claims referring to children’s development and health). Here, the scope of ‘function claims’ includes the ‘role of a nutrient or other substance in growth, development and the functions of the body, psychological and behavioural functions or without prejudice to Directive 96/8/EC, slimming or weight control or a reduction in the sense of hunger or an increase in the sense of satiety or to the reduction of the available energy from the diet’⁽³⁾. EFSA publishes scientific opinions on all types of claims⁽⁶⁾. Once accepted, (general) claims can be made without undergoing any further authorisation procedure. From the series of opinions published thus far⁽⁶⁾, much can be learned about EFSA’s line of thinking, for instance, on what type of effects are considered to be relevant to human health, what type of studies are considered to be pertinent to the health claim and how the quality of studies is judged. These issues are addressed in this paper.

Outside Europe, the regulatory landscape differs slightly. In the United States ‘reduction of disease risk’ claims have been allowed on certain foods since 1993. These foods contain components for which the FDA has accepted that for a correlation between nutrients or foods in the diet and risk of certain diseases, based on ‘the totality of publicly available scientific evidence’, there is substantial agreement amongst qualified experts that claims were supported by the evidence^(7–9).

The highest level of health claims in the United States are 'unqualified' health claims that confirm relationships between components in the diet and risk of disease or a health condition that are based on significant scientific agreement. Health claims can also be based on authoritative statements (resulting from a scientific body of the US Government or the National Academy of Science), and are permitted following notification to FDA and FDA's subsequent failure to object. Another category of health claims concerns so-called 'qualified' health claims that are used for describing developing relationships between components in the diet and disease. Such claims require qualifying language such as 'although there is scientific evidence supporting the claim, the evidence is not conclusive' and these claims require pre-market approval of the FDA. The FDA has published a list of approved unqualified claims⁽¹⁰⁾ and overviews of qualified health claims that have been approved⁽¹¹⁾ or denied⁽¹²⁾. In addition to approved FDA disease risk reduction claims, structure function claims are permitted on foods and dietary supplements. Structure function claims describe the role of a nutrient or dietary ingredient intended to affect normal structure or function in human subjects⁽¹³⁾. Structure function claims are not pre-approved by FDA, but must be truthful and must not be misleading. It is clearly recognised that, worldwide, health claims need to be scientifically substantiated.

Criteria for the scientific substantiation of health claims

Worldwide, there is broad consensus among scientists on the PASSCLAIM criteria for the scientific substantiation of health claims⁽¹⁾. These criteria aimed to provide a scientifically robust tool for evaluating the quality of data submitted in support of health claims for foods and to provide a standard against which the quality of existing evidence can be transparently graded.

In different regions of the world, regulatory bodies have published guidelines for manufacturers who want to apply for approval of certain health claims. These guidelines are broadly in accordance with the PASSCLAIM criteria. Current guidelines from international regulatory bodies can be summarised as follows:

- (1) Applications for the authorisation of health claims should adequately demonstrate that the health claim is based on and substantiated by generally accepted scientific evidence, by taking into account the totality of the available scientific data and by weighing the evidence.
- (2) A prerequisite of claim submission dossiers is a proper characterisation of the food or the food constituent, including composition, physical and chemical characteristics, manufacturing process, stability and (where applicable) bioavailability.
- (3) Human data are central for the substantiation of health claims on food products. Data from intervention studies are generally given more weight than observational data. Key considerations with respect to human studies include choice of an appropriate study group, study

design and execution, with appropriate controls and sufficient statistical power.

- (4) When markers of a target end point are used, these markers should be biologically and methodologically valid.
- (5) The target variable itself should change significantly and the change should be biologically meaningful for the target group. Also, the amount of food (constituent) should be consistent with the intended consumption pattern (i.e. the benefit should be achievable for the target population with a realistic use of the food product).

Experience with the scientific evaluation and use of health claims in recent years shows a clear need for further guidance on what would be required to prove the efficacy of food (constituents) and what evidence is needed to substantiate a health claim. Key issues to be considered are the following:

- (1) Well-designed randomised controlled trials (RCT) provide the most persuasive evidence of efficacy, allowing strong causal inferences. Other experimental studies, such as observational studies, identify associations between intake of food (constituent) and a beneficial effect on health, although it may be difficult to distinguish whether the observed difference in health status is due to differences in intake or to some other unrecognised (and often unmeasured) factor. Appropriate study design and statistical methods can be used to minimise the effects of such confounding variables. However, as will be discussed below, the use of RCT in evaluating clinical treatments and pharmaceuticals (evidence-based medicine) does not mean that this type of study is always the most appropriate approach for the evaluation of nutritional effects (evidence-based nutrition).
- (2) Relationships between dietary factors and diseases are likely to be extremely complex for both biological and behavioural reasons. Types and amounts of food eaten may be related to important non-dietary determinants of health and diseases, such as genetics, environment and lifestyle.
- (3) The identification and validation of relevant markers or marker patterns that reflect, or predict, potential benefits or risks relating to a target function in the body or risk factors for disease are important. There are, however, differences in approaches taken by various scientific bodies. It is important to assure the scientific robustness of markers for disease risk and their relevance to the key measure or target end point. There is also a need to focus research on identification and validation of such markers.
- (4) Studies should be performed according to current quality standards and results should be reproducible. Guidelines for the design, conduct and reporting of human intervention studies are described elsewhere⁽¹⁴⁾.
- (5) As considering the total body of evidence is the most appropriate way to judge substantiation of health claims⁽¹⁾, a scientific framework is needed to assess the strength, consistency, and biological plausibility of the evidence. As will be discussed below, the findings from several studies and their consistency should be weighed and integrated.

Building a scientific concept

Health claims require a systematic science-based evaluation of the strength of the evidence to support a food–health relationship. In order to build a scientific concept for this evaluation, several steps have to be taken, which include considering whether a claimed effect would imply a health benefit to a specific target group, selecting appropriate study types, target groups and markers or risk factors, considering the biological plausibility of the claimed effect and last but not the least weighing the totality of the available evidence.

Establishing benefit to health

Already at an early stage in the process of substantiating the health effect of a food (constituent), it has to be considered whether the effect to be claimed can reasonably be considered as ‘beneficial to health’, i.e. whether the intake of a food (constituent) results in a beneficial nutritional, physiological or psychological effect. Health is not typically defined within regulations and relates to the general condition of the body or mind; it is often considered within the context of the presence or absence of disease or impairments. The WHO definition of 1946 states ‘health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity’⁽¹⁵⁾. ‘Benefit’, on the other hand, has been defined, for example, in Europe by EFSA to be ‘the probability of a positive health effect and/or the probability of a reduction of an adverse health effect in an organism, system or (sub)population in reaction to exposure to an agent’⁽¹⁶⁾. With reference to European Union claims regulation, benefit to health relates to the weighing of the totality of the available scientific data. Depending on the outcome of this evaluation, conclusions may be drawn that the probability of a change in the dietary intake of a food (constituent) will result in a health benefit and/or a health outcome, including a change in a disease end point. This could be, for example, a fatal or non-fatal heart attack. The claimed effect must be clearly defined and the applicant should provide a rationale derived from the evidence as to why the claimed effect is considered nutritionally or physiologically beneficial, the relevance to human health and to what extent the effect has occurred. The use and definitions of modal verbs (e.g. may, might, can or will) or graphics could provide the opportunity to develop appropriate qualifying language or symbols to reflect the totality of the available data and the strength of the evidence.

A key word from the scientific assessment and regulatory perspectives is ‘probability’. Examples from Europe of physiological effects considered by EFSA to be ‘beneficial’ are the maintenance of normal levels/functions such as blood cholesterol/TAG, blood glucose, platelet aggregation, blood pressure, blood coagulation, energy-yielding metabolism and tooth mineralisation. With respect to children’s development and health, ‘beneficial’ has included normal visual development, cognitive development, growth and development of bone. Examples of risk factors accepted by EFSA include loss of bone mineral density in postmenopausal women, increased plaque acid (impacting on dental and oral health),

high blood cholesterol and an increase in potentially pathogenic intestinal microorganisms⁽⁶⁾.

In general, the extent of the causal relationship between the maintenance of normal and enhanced functions and human health should be demonstrated and should withstand scientific scrutiny, at least showing that if the function is not maintained within the normal range, human health can be affected. In addition, as well-being and quality of life may be related to physical states (‘health-related’), appropriate validated questionnaires may be used as part of claim substantiation.

Representative study group: the concept of health as a continuum

A beneficial health effect demonstrated in a certain study population can be extrapolated to another, or wider, population, given that the study population is representative of the target population. As most health claims on foods are intended for the ‘general public’, an important question concerns what constitutes a representative study group. Aiming to maintain health, claims target keeping an individual on the healthy side of boundaries between health and disease. As a continuum between health and disease often exists without a strict boundary, it would be reasonable to find within the ‘general population’ some individuals with slightly elevated blood pressure, slightly abnormal blood lipids, elevated markers of inflammation or mild liver impairment etc. Similarly, within the ‘healthy’ general population, a significant proportion of the population will be at risk of exhibiting certain risk factors for diabetes, heart disease or other chronic diseases and/or are overweight, but are otherwise going about their daily lives. To demonstrate certain beneficial effects, studies in a healthy population may not even be possible and, depending on the function being considered, evidence in slightly impaired individuals may represent an alternative.

The difficulty of clear demarcations between normal/healthy and unhealthy/diseased may be reflected in some examples. For instance, EFSA has accepted evidence for claims on reducing gastrointestinal discomfort (in the general population) evidenced in irritable bowel syndrome (i.e. in a patient population)⁽¹⁷⁾. In contrast, EFSA has not accepted evidence for claims relating to maintenance of normal joints (in the general population) evidenced by studies in those with osteoarthritis (i.e. in a patient population), since osteoarthritic cells/tissues may respond differently to intervention⁽¹⁷⁾. These opinions support that a case-by-case judgement would be necessary. However, as consistent decisions on the appropriateness of specific study groups are needed, it is strongly advisable to develop a clear rationale for decision making. We propose that the concept of health as a continuum (as depicted in Fig. 1) be used.

The proposed model in Fig. 1 considers that the general population, theoretically, comprises a range of individuals from those without symptoms of disease and who are in optimal health to those individuals who have an overt disease status. The health status of the majority of individuals lies on a continuum between these two extremes. In many cases (particularly those of chronic diseases), a ‘disease status’

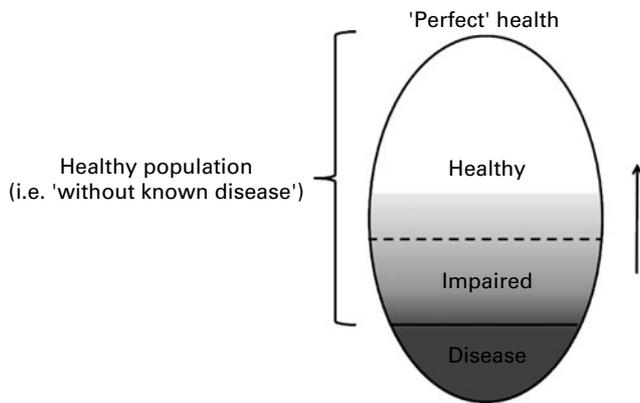


Fig. 1. Concept of 'health' as a continuum. The model is proposed to provide guidance on what effects/markers to consider when addressing physiological effects of food constituents on health. Health is considered as a continuum between perfect health (top) and disease (bottom) and reflects the general population. 'Perfect' health is a theoretical ideal that cannot be reached. Although usually the cut-off to disease is by convention set (unbroken horizontal line), an arbitrary cut-off between healthy and diseased (i.e. 'impaired') can be considered (broken horizontal line). Once the distribution of the effect/marker has been defined in the population, upward movement of the effect/marker (indicated by the upward arrow) would indicate lowering of the proportion of disease/impairment or maintaining health for longer and could be used to substantiate claims for the general population. For health claims addressing a health effect/marker for which a continuum (for most diseases) can be established, efficacy studies may use those with disease as a model rather than the target population only. The grading of 'grey' reflects the increase of symptoms.

develops as a continuum from a 'normal' to 'diseased' status. The proposed model considers such a continuum with health as a dynamic process. This includes sub-optimum functions that can be improved. For health targets without established markers, this model could be applied and evidence of protection against a decline in health status would be considered as reducing the risk of disease (and/or maintaining health). For most markers, a health benefit would be established if the value of the marker moves towards an optimum, i.e. towards the 'perfect' health part of the model. Equally, this could be applied to a distribution of markers. A claim would be substantiated by demonstrating a beneficial effect or maintaining the healthy status of individuals for longer time periods (for example, the reduced duration of an infection).

For health claims fitting this model, efficacy studies could use participants with disease as a model rather than the target population only.

In some cases, it may be difficult to demonstrate that healthy people will become healthier when consuming a specific food (constituent) with a putative health benefit. Overcoming this difficulty may be possible by challenging healthy individuals (e.g. exercise test) and observing recovery or by considering a marginally unhealthy group. For example, long-term maintenance of normal blood glucose concentrations is considered to be a beneficial physiological effect. The effect of a food (constituent) on postprandial hyperglycaemia in healthy participants could be considered. However, to consider longer-term effects on blood glucose concentrations only in studies in those with impaired glucose tolerance or those with type 2 diabetes mellitus, who are as

yet un-medicated, could generate appropriate evidence. Furthermore, if there is no reason to expect interference between the mechanisms of action between the drug(s) and the food (constituent), then studies in medicated people can also be used. Thus, populations in such studies used to support health claims may need to comprise participants who could be classed as healthy, marginally unhealthy or diseased. Although demonstrating health benefits for the general population is difficult, the use of well-defined subgroups may represent an acceptable alternative.

Selecting appropriate study types and designs

Health claims require the establishment or building of a scientific concept, followed by demonstration of effect often through human intervention studies. The number and type of studies required to demonstrate the claimed effect will vary according to the targeted benefit and already available knowledge/recognised scientific evidence. Also, the duration of studies should be defined in accordance with the intended beneficial effect and measures to be performed/markers to be followed and the targeted claim.

In nutrition research, usually RCT are given greater weight than (different types of) observational studies. Within the category of observational studies, findings in prospective cohort studies generally receive more weight than data from case-control and cross-sectional studies⁽¹⁸⁾. However, in the case of disease risk reduction claims, such as in CHD, a simple hierarchical approach to evidence on causal links cannot rely on RCT⁽¹⁹⁾. Given the complex nature of disease processes over decades, reliance on the use of relatively short RCT and risk factors/markers as primary sources of evidence for disease risk-reduction claims is questionable.

Observational studies can provide evidence of an association between consumption of a food (constituent) and health status rather than conclusive proof of cause and effect. However, properly designed and executed observational studies can provide a strong and consistent body of scientific evidence, which includes information on low-to-high quintiles of intake (i.e. an intake-effect relationship), a statistically significant measure of relative risk of developing the disease and true outcomes of the disease. For example, the US FDA has approved a health claim on the relation between intake of fruits and vegetables and a reduced risk of cancer⁽²⁰⁾.

Mechanistic studies, including both *in vitro* and *in vivo* (animal) studies, can provide important information to support the health relationship. However, complete elucidation of mechanisms should not be mandatory to support a claim (assuming the claim is not describing a mechanistic aspect). Indeed, for many pharmaceutical compounds (including well-known paracetamol), precise mechanisms of action are also not always known.

As various methodological problems and uncertainties are inherent to all study types, including RCT, a hierarchy of study types cannot be applied absolutely. Also, it should be noted that methodological soundness overrides any hierarchy in studies on human subjects, given that validity depends not only on the appropriateness of the study type but also on

the quality of its design, execution and analysis. The same difficulties occur with comparison of different studies within a meta-analysis⁽¹⁸⁾. As will be discussed later, the totality of the evidence needs to be assessed to determine the overall strength of the scientific convictions being proposed.

Challenges of meta-analyses. Meta-analytical approaches, with large sample sizes, are often superior to single-trial approaches. For instance, meta-analyses for the validation of surrogate end points have become a widely accepted and applied method in oncology research. Without multi-trial data, it is almost impossible to make any direct inference about the association between the diet and the surrogate and clinical end points, because one set of data cannot provide sufficient evidence of any association.

Doing a meta-analysis correctly demands expertise in both the method and the substance, and hence almost always requires collaboration between clinician(s) and an experienced statistician. Models based on random effects are very popular in meta-analyses, because they allow for inter-trial heterogeneity. A challenging problem in the implementation of a meta-analysis is to combine several studies in which similar medical outcomes or covariates are captured and considered. Furthermore, the meta-analysis needs to combine studies that demonstrate homogeneity for dietary composition, that are well characterised and that are for a similar specific claim. To obtain this, it is recommended that targeted meta-analyses be undertaken within the specific context of a specific claim.

It is impossible to say whether the results of a large RCT or those of a meta-analysis of many smaller studies are more likely to be close to the truth. Much depends on the details of both the research studies and the analyses. When both the trial and the meta-analysis appear to be of good quality, however, we would tend to believe the results of the large RCT.

Evidence-based nutrition v. evidence-based medicine. Although it is important to apply high standards of scientific investigation to the assessment of the impact of foods, the complex nature of nutrition means that assessing the impact of foods (constituents) may not be straightforward. Foods (constituents) are clearly not developed to alter specific body functions in order to prevent or treat diseases. Foods, as well as health claims, are aimed at the general/healthy population. Typically, in nutrition (evidence-based nutrition) there are more than one (or a few) principal end points or outcome measures, and the effects of food (constituents) may rarely be evaluated relative to its absence. In most cases, nutritional end points need to be measured over relatively long periods of time. Nutrients and other substances that contribute to nutritional or beneficial physiological effects tend to manifest themselves in small differences over longer periods of time. Nutrients work together, rather than in isolation, and often their effects will not develop when intakes of other dietary components are suboptimal. There is, in effect, rarely a nutrient-free state against which the nutrient effects can be compared. The dilemmas of focusing on pharmaceutical approaches to evidence-based nutrition are highlighted by Heaney⁽²¹⁾ and Blumberg *et al.*⁽¹⁸⁾.

The reliance on RCT to assess the impact of food (constituents) fails to address the limitations of this pharmaceutical approach for nutrition and may explain, at least in part, the heterogeneity of results from different research centres and investigators and the different sources of evidence. For example, in certain nutrition studies, controls (e.g. placebo groups) may be difficult to define^(13,18). Thus, when considering approaches to determine the efficacy of foods (constituents), it is important to differentiate between nutritional practice (evidence-based nutrition) and drug practice (evidence-based medicine).

Health claims relating to children. For claims relating to consumption in children, evidence for substantiating the claim should be generated in the target age group. However, cases in which the physiology of a certain function of a younger age group can be demonstrated to be equally applicable to an older group, evidence generated in that older age group would be sufficient. As with adults, studies in children must be of high quality, but it must be realised that possibility for standardisation of studies may be reduced when studying children (as compared to adults) owing to ethical constraints, e.g. blood sampling. Consideration may also be needed when assessing study quality (*v.* feasibility) of such scientific evidence. For example, cross-over trials may not be possible due to physiological changes related to growth/development in the participants.

Choice of risk factors/markers

Valid (bio)markers are essential for substantiating claims or messages using a combination of intervention and observational studies together with appropriate analyses. However, it is important to differentiate between markers for disease (with regard to drugs) and markers for health/risk of disease (with regard to foods). For the substantiation of function claims, a beneficial nutritional, physiological, psychological (e.g. cognitive/mental performance) or behavioural effect needs to be demonstrated. A marker (or set of markers) for a function is a measurable parameter that is indicative of the state of a particular function and thus helps to determine the effect of a food on a function and the state of health of an individual. According to PASSCLAIM, markers should be biologically valid, in that they have a known relationship to the final clinical outcome, and their variability within the target population must be known⁽¹⁾. Methodologies to measure the marker would need to be developed across a range of values or at a threshold value corresponding to the healthy and normal function of organs/tissues. For example, flow-mediated dilation is a recognised indicator of blood vessel functionality and can predict whether there is (or is not) dysfunction.

It is relevant to note that the EFSA definition of a 'reduction of disease risk' claim implicates that it is not a reduction of the risk for developing a disease that can be claimed, but only the reduction of a risk factor for the disease. Defining a suitable risk factor for disease may be complicated. According to the National Institutes of Health⁽²²⁾, a risk factor increases the likelihood of development of a disease rather than predicting it.

According to EFSA⁽²³⁾, a 'risk factor is a factor associated with the risk of a disease that may serve as a predictor for the development of that disease'. This definition is rather general. Most risk factors are correlational and are not necessarily causal. The relationship of a risk factor to the development of a disease should be biologically plausible. Furthermore, decreasing a risk factor should also be associated with a reduction of the risk for the disease. In certain cases, a surrogate end point (a biomarker intended to substitute for a clinical end point that should predict clinical benefit or harm or lack of both) can be used. There are also cases in which a reduction in the incidence of specific diseases can be used to support a more general claim, which is related to the reduction of a risk factor. For example, if it is adequately demonstrated that consumption of a food (constituent) decreases specific gastrointestinal or respiratory tract infections, this can help to define a risk factor that correlates with 'resistance to infections' or 'immune function'.

The choice of appropriate risk factors and biomarkers depends on the objective of a particular study. Markers and/or risk factors can be used at different biological levels (i.e. at the cellular, organ, individual participants or at a population level). The choice of biomarker(s) should be validated with agreement amongst independent experts in the field based on previous research using different scientific approaches (*in vitro*, animal, intervention studies in healthy/pre-symptomatic/unhealthy volunteers or cohort studies).

Established physiological risk factors, all of which are currently regarded as disease related, include the following: raised blood pressure, raised plasma insulin, raised fasting glucose, raised plasma total cholesterol and LDL, raised LDL:HDL ratio or loss of bone mineral density. Emerging risk factors include ghrelin levels (for obesity), faecal calprotectin (for inflammatory bowel disease) and adenomatous polyps (for colorectal cancer). These risk factors are based on pharmacological and disease responses, and are not necessarily indicators of the normal, robust, homeostatic mechanisms and state of health and resistance to disease. When other risk factors are proposed, the applicant is required to provide scientific justification. As stated later in the final recommendations of this paper, validating/establishing such risk factors should be a research priority for industry and academia.

In certain cases, a set of markers is more convincing than a single biomarker. One example is the guidelines for the global assessment of symptom relief in irritable bowel syndrome trials⁽²⁴⁾. In this case, the composite scores of frequency of digestive complaints is the sum of frequency of four individual digestive complaints (i.e. abdominal pain/discomfort, bloating, flatulence/passage of gas and borborygmi/rumbling stomach) and these can be evaluated using a five-point scale ranging from 0 ('never') to 4 ('every day of the week'). Key outcomes can also be based on questionnaires assessing either a specific condition or making a health or nutritional assessment. In recent years, the use of questionnaires has gained greater recognition by experts in the field of digestive health. A questionnaire must be developed for use in the population in which the effect of a food (consistent) will be demonstrated. The method used for the development of the

questionnaire must follow general recommendations for the development of participant-reported outcomes and accepted methods in the area of research. The validation of the questionnaire must comprise the validity of the concept measured in the target population, the tool of measuring (e.g. Likert scale or visual analog scale), the reference period of the measure (e.g. daily, weekly, monthly), as well as the method of measuring (e.g. method of presenting questions, instruction given to participants answering the questions). The sensitivity of the questionnaire should be supported by its ability to discriminate populations with different levels of the concept measured. The description of the factors that may influence the measures (i.e. confounding factors) is an important point for the use of the questionnaire in intervention trials.

However, in several areas of research, no validated physiological markers exist. With the exception of well-established risk factors, in Europe a case-by-case approach is taken by EFSA to assess the extent to which the reduction of a risk factor is beneficial in the context of a given disease risk reduction claim⁽¹⁷⁾. Reduction of risk factors/markers from human intervention studies appear to be given greatest priority. For example, EFSA⁽¹⁷⁾ considered that 'dietary behaviour' (e.g. diets with low content of a specific category of foods) would not be acceptable as a risk factor in this context (of 'reduction of disease risk') as the beneficial alteration of the risk factor (increased consumption of a specific category of foods) would not be considered as a beneficial physiological effect, as required by the Regulation (1924/2006). However, modifiable behavioural risk factors, such as diet, are relevant to underpin health claims, especially when the claim has the potential to enhance consumers' knowledge of healthy eating patterns and when the health claim complements well-established dietary recommendations⁽²³⁾.

For the substantiation of claims in children, apart from growth and development, food (constituents) may also aim to affect acute and/or longer-term functions. For example, prolonged modulation of nutrients in early life may influence brain development in a manner that permanently affects visual function or intelligence quotient of children. On the other hand, cognitive functions such as reaction time, memory and information processing speed, attention and 'normal activity' may be variable and influenced acutely by nutrients and other environmental factors. If food (constituents) acutely modulated such functions, then their effect would be related to the acute intake rather than sustained intakes and permanent stimulation. Putative long-term effects such as improved academic achievement could also be considered, although influences from environment and lifestyle variables are difficult to control.

An appropriate statistical procedure needs to be used when multiple exposures are examined using a 'multiple comparison' approach. The *P* value used for statistical significance should be adjusted according to the number of variables examined. These techniques generally require a stronger level of evidence to be observed in order for an individual comparison to be deemed 'statistically significant', so as to compensate for the number of inferences being made. Multiple testing corrections refer to re-calculating probabilities

obtained from a statistical test that was repeated multiple times. The use of *post hoc* analyses or looking at the data (after the study has concluded) for patterns that were not specified *a priori* also has some limitations. Each time a pattern in the data is being considered, a statistical test is effectively being performed. This greatly inflates the total number of statistical tests used and necessitates the use of multiple testing procedures to compensate for this. Results of *post hoc* analysis should be explicitly labelled as such in reports and publications.

Biomarker patterns in relation to homeostatic adaptability. There are relatively few validated biomarkers and risk factors applied to foods. Therefore, the identification of further relevant marker(s) to measure food functionality in the human body is one of the most important challenges within nutrition research today. The key challenges facing researchers are: first, identifying the link between a marker (or a set of markers) and a function or between marker(s) and the risk of a disease; second, it is important to identify a method to measure such marker(s); third, to observe the impact of a particular food on such marker(s) and hence its impact on health. On the basis of the principle that nutrition primarily aims to maintain or possibly improve health, new methods and models are currently being developed that better take into account the complexity and balance of homeostatic mechanisms. These models are based on dynamic processes instead of single end points. Recent advances in genomics and systems biology enable researchers to measure and model biomarker profiles and to translate these into dynamic processes. On the basis of the classical principles of

homeostasis and biological evolution, it is proposed that the term 'health' be defined as 'the ability to adapt' to internal and external stimuli. In case of chronic pathology or slow-developing pathologies, it can be said that there is an adaptation, as the individual can live with it for very long period of time, even without medication. However, this adaptation does not mean this individual would be considered to be a healthy person. Thus, as illustrated in Fig. 2, the individual's homeostasis acts to maintain balance within biological processes and is reflected by clusters of functional biomarkers that are kept within a certain range.

Clusters of biomarkers that reflect essential processes, such as inflammation or oxidative stress, can be used to construct a theoretical multi-dimensional 'health space'⁽²⁵⁾. These models can be used to illustrate effects of nutritional intervention on homeostatic balances in individuals. A biomarker approach may be used to detect early signs of homeostatic disturbance, as observed, for example, at onset of disease. Indeed, adaptive states, such as a chronically increased inflammatory status, clusters of cardiovascular risk factors and/or specific changes in metabolic fluxes, may be used as indicators of suboptimal health well before there is any clinical sign of disease. It is recognised that large inter-individual differences in 'normal' biological processes values exist, which give rise to an added complexity.

A broader and probably more predictive indication of health status is obtained by measuring the robustness (adaptability) of the processes of homeostasis in an individual. It is well-accepted that when an organism is challenged (i.e. when its system is disturbed) various compensatory

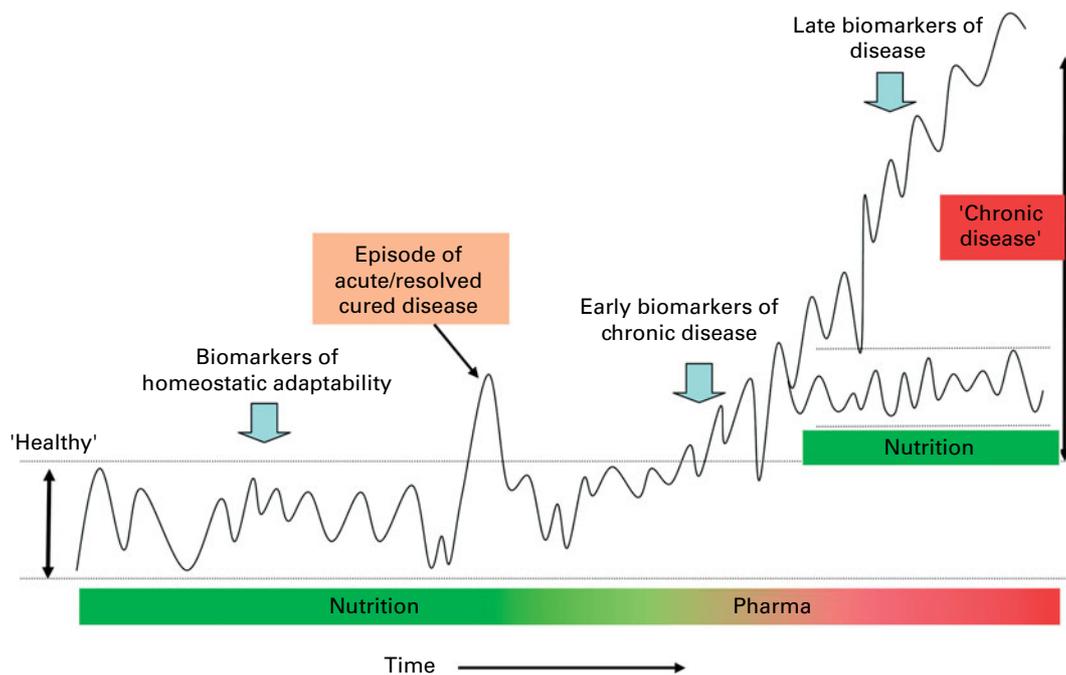


Fig. 2. Biomarker patterns in relation to homeostatic adaptability. Schematic depiction of the concept of physiological balance and the significance of biomarker patterns for various stages of development in time from normality (homeostasis), via dysfunction, to chronic disease. An organism maintains homeostasis for as long as possible by changes in its metabolic pathway dynamics. Nutrition aims to support this homeostasis. Chronic disease develops when an organism (individual) is no longer able to maintain homeostatic processes within a certain limit and may require intervention. A disease process can either further deteriorate or stabilise at a new homeostatic state.

mechanisms are used so that homeostasis is maintained for as long as possible. Challenge tests may be used in nutrition and health research to measure this adaptability, and such tests include variations of standardised oral glucose and lipid tolerance tests, organ function tests, infection challenge tests, exercise stress and psychological stress challenges. When combined with newer bioanalytical and statistical tools, such standardised tests (such as the oral glucose tolerance test) may be greatly enhanced, making them particularly useful to test health-improving effects of nutritional products.

Biological plausibility

In the context of nutrition and health claims, biological plausibility can be defined as the probability that, or the extent to which, a causal effect is demonstrated between the food (constituent) and the claimed physiological or psychological effect in human subjects, which is consistent with existing biological knowledge. From a regulatory point of view, the weight that is given to biological plausibility in claims substantiation is not consistent across the globe. EFSA has indicated that a rationale or evidence on biological plausibility of the claimed effect should be provided to support the substantiation of the claim⁽²⁶⁾. The US FDA, however, states that animal and *in vitro* studies can be used to generate hypotheses, investigate biological plausibility of hypotheses or explore mechanism(s) of action of a specific food (constituent) through controlled animal diets; however, these studies do not provide information from which scientific conclusions can be drawn⁽⁹⁾. Also, Health Canada indicates that, if desired, non-human studies may be used to support the discussion on biological plausibility and that this is optional⁽²⁷⁾. Biological plausibility is one of the most difficult causation criteria. In nutrition science, the biological evidence is collected from animal models, *in vitro* cell systems and human metabolic and intervention studies. There is as yet no consensus on the relative importance of each of these types of evidence, and decisions on usefulness tend to be subjective at present, especially with regard to causal inference. Biological plausibility is closely related to understanding the mechanism of action for the food (constituent) of interest. Thus, when assessing biological plausibility, the existence and/or relevance of possible multiple biological functions of the food (constituent) needs to be considered. Furthermore, with the complex inter-relationships of nutrients in the diet coupled with the potential for metabolic nutrient–nutrient interactions, a simplistic approach can be misleading. Also, the co-existence of constituents in the same foods and in associated foods provides an opportunity for multiple mechanisms. Consequently, biological plausibility has to be established on a case-by-case basis.

Weighing the totality of the evidence

Health claims require a high standard of evidence. PASSCLAIM established a robust standard with which it is possible to compare the quality of state of the art nutritional scientific data submitted in support of health claims and provide a basis for the harmonisation of the scientific evaluation and approval

of such claims⁽¹⁾. Also, the requirement for assessing the totality of the scientific data and weighing the evidence is built into current legislative regulations. The assessment of each specific food–health relationship that forms the basis of the claim is therefore based on a scientific judgement on the extent to which a cause and effect relationship is established by taking into account the nature and quality of different sources of evidence. In each case, the evidence is weighed with respect to its overall strength, consistency and biological plausibility (i.e. likelihood), but currently a grade of evidence is not assigned. In Europe, for instance, the outcome of each assessment has one of three conclusions: (1) a cause and effect relationship has been established between the consumption of the food (constituent) and the claimed effect (i.e. the claim is substantiated by generally accepted scientific evidence); (2) the evidence provided is insufficient to establish a cause and effect relationship between the consumption of the food (constituent) and the claimed effect (i.e. cause and effect is not conclusive because the evidence is emerging and/or conflicting, and the claim is thus not substantiated by ‘generally accepted scientific evidence’); (3) a cause and effect relationship has not been established between the consumption of the food (constituent) and the claimed effect (i.e. there is no or, at most, limited scientific evidence, and thus the claim is not supported by ‘generally accepted scientific evidence’)⁽¹⁷⁾. The authors acknowledge the importance of understanding what ‘generally accepted scientific evidence’ entails.

PASSCLAIM considered both the evaluation of the totality of the data and weighing of the evidence to be important in view of different interpretations of conflicting evidence and the potential variation in quality amongst individual studies. Not all research has been, or will be, carried out to the highest standard, or even to a common standard. This may, in part, be due to the complexities of research in human subjects and also because data in support of a claim may have been taken from studies that had a different primary objective. Despite potential limitations in the research base, there may be complementarity between individually incomplete studies that support an assessment of the totality of the evidence to substantiate a claim, for example, using a meta-analysis. Conversely, a review of all the studies taken together may reveal evidential inconsistencies that are not apparent from the review of a single study in isolation⁽¹⁾. PASSCLAIM also stated that any template needs to be applied intelligently and sensitively to the existing and potential claims on a case-by-case basis, with respect to both gaps in knowledge and to the development of new knowledge. Although PASSCLAIM provided a scientific framework to facilitate the assessment of scientific support for claims on foods, it did not specifically address weighing of the evidence. However, it was later emphasised that the evaluation process should be transparent and that the grading of evidence into categories, including ‘convincing’, ‘probable’, ‘possible’ and ‘insufficient’, could be considered useful in scientific evaluations and to monitor the development of the scientific substantiation for the claim⁽²⁸⁾.

The development of a scientific framework for weighing the totality of the available data and the determination of the extent to which a cause and effect relationship is demonstrated are both scientifically justified. However, there is currently no consensus about how the beneficial associations between foods (constituents) and health can be tested and established, and indeed whether the requirement for conclusive evidence of cause and effect is proportionate and achievable in nutrition science. Key questions relate to what constitutes the totality of the evidence and by what means it should be developed and weighed. Although some guidelines have been provided, there is still a need for a clear framework for the assessment of the strength of the evidence, otherwise applicants for a health claim will not be clear on the research programs they will need to construct to substantiate a claim. It is necessary to have a transparent framework for commenting on the nature and quality of the totality of the data and for weighing the evidence in order to allow independent experts to judge about the scientific evidence of a health claim submitted by an applicant.

(Inter)national organisations have used various systems to assess the level of evidence from different types of studies. One common approach is the distinction between different levels of evidence. This classification of the evidence into

categories (e.g. 'convincing', 'probable', 'possible' and 'insufficient') has proven to be a useful tool in scientific evaluations. For instance, the WHO⁽²⁹⁾ and the World Cancer Research Fund⁽³⁰⁾ have published comprehensive and rigorous evaluations of the strength and consistency of evidence for the relationship between certain nutrition factors and different chronic diseases, with judgments characterised as being 'convincing' or 'probable' considered strong enough to justify population goals and personal recommendations^(29,30). In the recent EFSA scientific opinion on establishing food-based dietary guidelines⁽³¹⁾, the identification of diet–health relationships was described using the same terminology, namely, convincing evidence, probable evidence, possible evidence and insufficient evidence. Likewise, the EFSA consultation paper on guidance on human health risk–benefit assessment of foods⁽¹⁶⁾ defines 'benefit' as the probability of positive health effects and/or the probability of a reduction of an adverse health effect. Other researchers⁽³²⁾ have proposed similar approaches for assessing the strength of the evidence and identifying the criteria for the use of the terms strong, moderate or weak. Although the classification could be criticised for being arbitrary, this framework illustrates that it is possible to assess the extent of the evidence of causation and to compare the consistency of relative risks

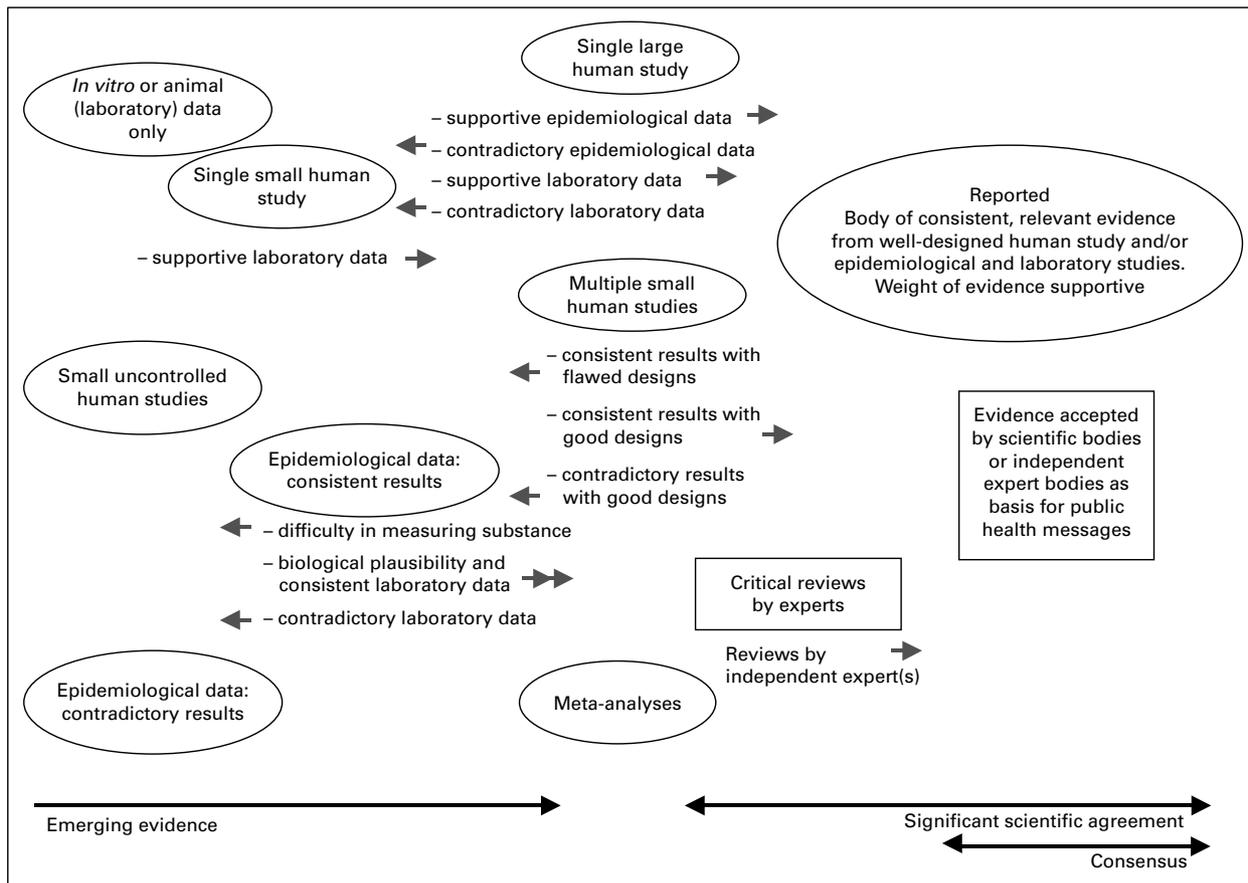


Fig. 3. Graphical representation for weighing of the evidence on a case-by-case basis in support of a health claim. †Adapted from Richardson *et al.*⁽³³⁾. The arrows reflect the fact that the totality of the evidence is made up of different sources of scientific data, and that each health relationship and claim must be assessed on a case-by-case basis to reflect the strength, consistency and coherence of the information. The graphical representation also reflects the scientific method, in that individual study results can be inconsistent, but as the science evolves, consensus may evolve, which allows the balance of probabilities for the scientific link between a food (constituent) and a health benefit to be assessed⁽³³⁾.

from cohort studies with outcomes from RCT. The findings support the strategy of investigating dietary patterns in cohort studies and RCT, especially for common and complex multi-causal chronic diseases such as CHD.

Clearly, the concepts developed by PASSCLAIM⁽¹⁾, WHO⁽²⁹⁾, the World Cancer Research Fund⁽³⁰⁾ and EFSA⁽³¹⁾ could be used further to underpin the assessment of the totality of the available data and, in particular, the weight of the evidence such as that illustrated in Fig. 3 on a case-by-case basis.

Conclusions and key recommendations

The identification of a suitable scientific framework for the weighing of evidence is now critical to embrace 'state of the art' nutrition science, to stimulate future academic research, to promote product innovation and to communicate accurate and truthful nutrition and health messages to the public.

In conclusion, the substantiation of a health claim needs to take place on a case-by-case basis. As a first step to a substantiation process, a strategy is needed to ensure scientific consensus, which includes input from independent scientific experts and, if possible, regulatory authorities. Such a strategy includes elements such as the benefit (in the targeted claim) to health, considerations concerning what constitutes the 'healthy population model', selection of the appropriate study target groups' considerations and decisions on the extent to which the mechanism of action will be established and which (bio)markers and risk factors will be used and (if necessary) validated, type, design and number of studies enabling demonstration of reproducibility of the claimed effect. All these elements have been considered in this paper. In executing the substantiation strategy, any deviations should be clearly explained and documented. Where such a deviation is significant, it is recommended that scientific consensus is re-established. The chosen strategy and its execution should be subsequently included in the scientific dossier that is submitted to substantiate the targeted health claim.

In addition, it is recommended that:

- (1) Further discussion is needed on the basis for accepting whether a demonstrated effect can be considered as beneficial to health.
- (2) A suitable scientific framework should be agreed that addresses the relationship between intervention and observational studies, taking account of characteristics of the food constituent (or nutrient), quality of the studies, appropriateness of study populations, confounding (in observational studies) and design of the intervention studies.
- (3) A suitable scientific framework for the weighing of evidence should be agreed.
- (4) Validating/establishing risk factors should be adopted as a future research priority. Expert groups should be convened to provide consensus on the level of acceptability of emerging risk factors.
- (5) A continuum of health approach should be applied where applicable.

- (6) Models based on measuring the adaptability of homeostatic processes should be further developed and evaluated, preferably by international research consortia in which industry and academic groups work together.

Acknowledgements

The authors thank participants of the workshop 'Beyond PASSCLAIM—Guidance to substantiate health claims on foods' held in Nice, France from 14 to 16 December 2009⁽²⁾. This workshop brought together over seventy experts from industry, academia and public bodies to discuss guidelines to establish beneficial effects of functional foods, and outputs of the workshop have been incorporated into this paper. The work was commissioned by the Functional Foods Task Force of the European branch of the International Life Sciences Institute (ILSI Europe). Industry members of this task force (in 2009 and 2010) are Abbott Nutrition, Barilla G. & R. Fratelli, Bayer CropScience BioScience, Bionov, Cadbury, Cargill, Coca-Cola Europe, Colloides Naturels International, CSM, Danisco, Danone, Dow Europe, DSM, FrieslandCampina, Frutarom, International Nutrition Company—INC, Kellogg Europe, Kraft Foods, La Morella Nuts, Mars, Martek Biosciences Corporation, McNeil Nutritionals, Monsanto Europe, Naturex, Nestlé, PepsiCo International, Pfizer, Puleva Biotech, Red Bull, Rudolf Wild KG, Soremartec Italia—Ferrero Group, Südzucker/BENEOL Group, Syral, Tate & Lyle, Ülker Bisküvi, Unilever, Valio and Yakult Europe. The opinions expressed herein are those of the authors and do not necessarily represent the views of ILSI Europe nor those of its member companies. G. W. M. is employed by Unilever, M. S. was employed by ILSI Europe at the time of the publication and G. T. is employed by Danone. Both D. P. R. and M. S.-W. are specialist consultants in food/nutrition science, advising research institutes, (non)governmental agencies and food industry. For those experts affiliated with academic or non-industrial institutions, ILSI Europe covered the expenses related to their participation in expert group meetings and the workshop held in Nice, and an honorarium has been provided.

References

1. Aggett PJ, Antoine J-M, Asp N-G, *et al.* (2005) Process for the assessment of scientific support for claims on foods (PASSCLAIM). *Eur J Nutr* **44**, Suppl. 1, 1–30.
2. Aggett PJ, Antoine J-M, de Vries J, *et al.* (2010) Beyond PASSCLAIM – guidance to substantiate health claims on foods. Summary report of workshop held in December 2009 in Nice, France. Organised by the ILSI Europe Functional Foods Task Group. <http://www.ilsilife.org/Europe/Publications/Beyond%20PASSCLAIM%20Report%202010.pdf> (accessed October 2010).
3. European Parliament and Council (2007) Corrigendum to Regulation (EC) no. 1924/2006 of the European Parliament and of the Council of 20th December 2006 on nutrition and health claims made on foods. *OJ L12*, 3–18. http://eur-lex.europa.eu/LexUriServ/site/en/oj/2007/l_012/l_01220070118en00030018.pdf (accessed October 2010).

4. US FDA (2003) Claims that can be made for conventional foods and dietary supplements. <http://www.fda.gov/Food/LabelingNutrition/LabelClaims/ucm111447.htm>. (accessed October 2010).
5. Codex Alimentarius (1997/amended 2009) Nutrition and health claims (CAC/GL 23-1997). http://www.codexalimentarius.net/web/more_info.jsp?id_sta=351 (accessed October 2010).
6. EFSA Panel on Dietetic Products, Nutrition and Allergies (2010) Scientific Opinions on the substantiation of health claims. <http://www.efsa.europa.eu/en/nda/ndaclaims.htm> (accessed October 2010).
7. US FDA/CFSAN (1998) Guidance for industry: notification of a health claim or nutrient content claim based on an authoritative statement of a scientific body. <http://www.fda.gov/Food/GuidanceComplianceRegulatoryInformation/GuidanceDocuments/FoodLabelingNutrition/ucm056975.htm>
8. US FDA/CFSAN (2007) Qualified health claims subject to enforcement discretion. <http://www.fda.gov/Food/LabelingNutrition/LabelClaims/QualifiedHealthClaims/ucm073992.htm> (accessed October 2010).
9. US FDA (2009) Guidance for industry: evidence-based review system for the scientific evaluation of health claims. <http://www.fda.gov/Food/GuidanceComplianceRegulatoryInformation/GuidanceDocuments/FoodLabelingNutrition/ucm073332.htm> (accessed October 2010).
10. US FDA (2009) Health claims meeting significant scientific agreement (SSA). <http://www.fda.gov/Food/LabelingNutrition/LabelClaims/HealthClaimsMeetingSignificantScientificAgreementSSA/default.htm> (accessed October 2010).
11. US FDA (2009) Qualified health claims: letters of enforcement discretion. <http://www.fda.gov/Food/LabelingNutrition/LabelClaims/QualifiedHealthClaims/ucm072756.htm> (accessed October 2010).
12. US FDA (2009) Letters of denial. <http://www.fda.gov/Food/LabelingNutrition/LabelClaims/QualifiedHealthClaims/ucm072751.htm> (accessed October 2010).
13. US FDA (2008) Food labeling guide. Guidance for industry. <http://www.fda.gov/Food/GuidanceComplianceRegulatoryInformation/GuidanceDocuments/FoodLabelingNutrition/FoodLabelingGuide/default.htm>
14. Welch RW, Woodside JV, Antoine J-M, *et al.* Guidelines for the design, conduct and reporting of human intervention studies to evaluate the health benefits of foods. *Br J Nutr* (accepted).
15. World Health Organisation (1946) Preamble to the Constitution of the World Health Organization as Adopted by the International Health Conference, New York, 19–22 June, 1946.
16. EFSA (2010) Guidance on human health risk-benefit assessment of food. *EFSA J* **8**, 1673. <http://www.efsa.europa.eu/en/scdocs/doc/1673.pdf> (accessed October 2010).
17. EFSA Panel on Dietetic Products, Nutrition and Allergies (2011) General guidance for stakeholders on the evaluation of Article 13.1, 13.5 and 14 health claims. *EFSA J* **9**, 2135. Published online: <http://www.efsa.europa.eu/en/efsajournal/doc/2135.pdf> (accessed June 2011).
18. Blumberg J, Heaney RP, Huncharek M, *et al.* (2010) Evidence-based criteria in the nutritional context. *Nutr Rev* **68**, 478–484.
19. Wiseman M (2008) The second World Cancer Research Fund/American Institute for Cancer Research expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. *Proc Nutr Soc* **67**, 253–256.
20. Code of Federal Regulations (2003) Title 21 – Food and Drugs, Volume 2, Chapter I – Food and Drug Administration, Department of Health and Human Services, Part 101 – Food Labeling, Sec. 101.78 Health Claims: Fruits and Vegetables and Cancer, pp. 134–136.
21. Heaney RP (2008) Nutrients, endpoints and the problem of proof. *J Nutr* **138**, 1591–1595.
22. NIH Definition Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* **69**, 89–95.
23. EFSA Panel on Dietetic Products, Nutrition and Allergies (2011) Guidance on the scientific requirements for health claims related to gut and immune function. *EFSA J* **9**, 1984. Published online: <http://www.efsa.europa.eu/en/efsajournal/doc/1984.pdf> (accessed June 2011).
24. Irvine EJ, Whitehead WE, Chey WD, *et al.* (2006) Design of treatment trials for functional gastrointestinal disorders. *Gastroenterol* **130**, 1538–1551.
25. Van Ommen B, Fairweather-Tait S, Freidig A, *et al.* (2008) A network biology model of micronutrient related health. *Br J Nutr* **99**, Suppl. 3, S72–S80.
26. EFSA Panel on Dietetic Products, Nutrition and Allergies (2009) Frequently Asked Questions (FAQ) related to the assessment of Article 14 and 13.5 health claims applications on request of EFSA. *EFSA J* **7**, 1339 (8 pp.).
27. Bureau of Nutritional Sciences (2009) Food Directorate, Health Products and Food Branch. Guidance document for preparing a submission for food health claims. Health Canada. http://www.hc-sc.gc.ca/fn-an/alt_formats/hpfb-dgpsa/pdf/legislation/health-claims_guidance-orientation_all_egations-sante-eng.pdf (accessed October 2010).
28. Asp NG & Bryngelsson S (2008) Health claims in Europe: new legislation and PASSCLAIM for substantiation. *J Nutr* **138**, 1210S–1215S.
29. World Health Organisation (2003) *Diet, Nutrition and the Prevention of Chronic Diseases: Report of a Joint FAO/WHO Expert Consultation*. WHO Technical Report Series 916. World Health Organisation Geneva.
30. World Cancer Research Fund International/American Institute for Cancer Research Expert Report (2007) Food, nutrition, physical activity and the prevention of cancer: a global perspective. <http://www.dietandcancerreport.org/> (accessed October 2010).
31. EFSA (2010) Scientific opinion on establishing food-based dietary guidelines. *EFSA J* **8**, 1460. <http://www.efsa.europa.eu/en/scdocs/doc/1460.pdf> (accessed October 2010).
32. Mente A, de Koning L, Shannon H, *et al.* (2009) A systematic review of the evidence supporting a causal link between dietary factors and coronary heart disease. *Arch Intern Med* **169**, 659–669.
33. Richardson DP, Affertsholt T, Asp N-G, *et al.* (2003) PASSCLAIM – synthesis and review of existing processes. *Eur J Nutr* **42**, 96–111.