The more remote markers are from the endpoint, the less specific and more attenuated and subject to confounding variables they become. Conversely, they become more specific and quantitatively related the closer they are to the endpoint in question. The characterisation of the mechanisms and pathways leading to outcomes would refine the identification of markers and inform how they may be selected. The generation of this knowledge, including approaches based on genomics and post-genomic molecular biology, underpins the biological and physiological validity of markers (see criteria 4 and 5), is fundamental to advances in nutrition, and integral to the development of foods with claims (nutrient function claims, enhanced function claims and reduced risk of disease claims) (Fig. 3).

All markers, irrespective of whether they are biochemical, physiological or behavioural in nature, should be valid (see criterion 4).

In some cases an individual marker may not provide sufficiently robust support for the desired claim. It may be that a combination of several relevant but not necessarily closely related markers can be used to justify the claim. This approach would need biological and statistical evaluation and an understanding of the independent strengths of association and the overall probability that their combined use strengthens the justification of the claim.

Criterion 4. Markers should be:
- **biologically valid in that they have a known relationship to the outcome and their variability within the target population is known;**
- **methodologically valid with respect to their analytical characteristics.**

There should be evidence that any particular marker reflects a meaningful biological effect and can be reliably and reproducibly measured. The validity of a marker comprises two aspects: 1) the biological validity and 2) the technical or methodological validity. Whereas the biological validity is common to all laboratories, the methodological validity needs to be established for each laboratory.

Biological validity. Biological validity concerns the extent to which a marker reflects a certain health outcome of interest and the process leading to it. It is not dependent on the technical competence of any individual laboratory. The biological validity of a marker derives from its relationship to the biological processes leading to the health effect and requires that the marker changes in line with a changing event or circumstances (for example the consumption of a particular food). In addition to insight into the biological process, it is necessary to have knowledge of the sensitivity and specificity of the marker for the health effect (see Annex 1, [34, 35]). As is noted above (criterion 3), a marker is not the same outcome as the health endpoint. The existence of an association between a marker and a disease risk does not necessarily mean that changing the variable changes the disease risk. Such modification can be effective only if the relationship is causal, and if effects already induced are reversible [36]. Hence the appropriateness of a marker needs to be considered on a case-by-case basis. It might be that a single marker does not meet all the criteria required for complete substantiation of a health effect. The marker may nevertheless contribute usefully to the totality of the evidence (see criterion 3).

Methodological validity and quality control. Any laboratory performing measurements should be competent to perform the measurements and to certify that the values produced can be trusted – that is, the method is technically valid in its performance by that laboratory. Study requirements for documentation and control, such as Good Laboratory Practice (GLP), Good Clinical Practice (GCP) and Good Epidemiological Practice (GEP),

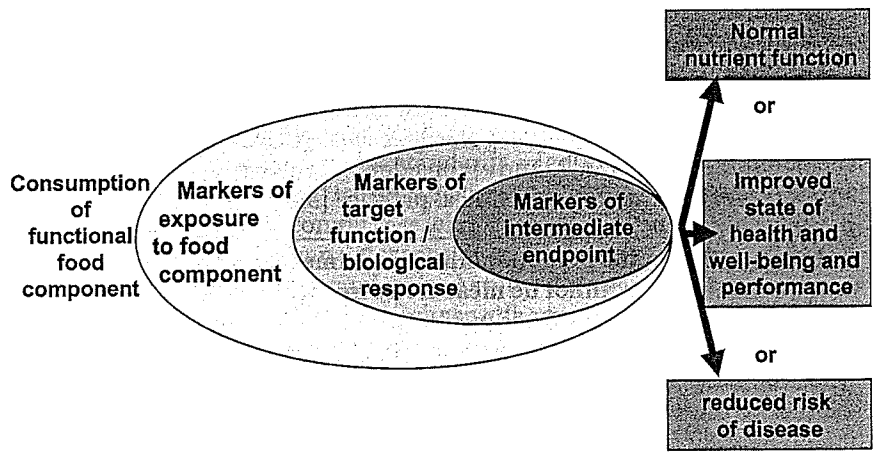

Fig. 3 PASSCLAIM classification of markers relevant to health claims

should not be confused with technical requirements prior to running analyses. As concerns the latter, Quality Control (QC) is important for claims in terms of technical validation of measurements and encompasses aspects such as accuracy, precision, repeatability, reproducibility and linear and dynamic range (see Annex 2). Requirements for these can be found on web-sites from chemical societies and national and international committees for analytical validation (for example www.fasor.com/iso25/, www.aoac.org, www.nmkl.org, www.ich.org). During method development, data on validity can be collected and compiled in a test method dossier, which is unique to each laboratory. After method validation, routine analyses can be performed. For these, quality control is typically performed by running concurrent control samples, and checking the actual results versus means and their standard deviation.

The total variability in the measurement of any parameter of interest is a combination of the biological variability and the methodological variability. The best results in a study can be obtained by having insight into the biological and the methodological validity at the design stage of the study.

Whereas, historically, research using markers has been done in a reductionist way (that is, by using one or only a few markers simultaneously), genomic and post-genomic molecular biology can perhaps generate a more integrated approach including molecular and whole body studies to establish claims. Even so the requirements for biological and methodological validity will remain.

Criterion 5. Within a study the target variable should change in a statistically significant way and the change should be biologically meaningful for the target group consistent with the claim to be supported.

This criterion reflects the importance of both the statistical significance and the biological meaningfulness of an effect.

At the level of statistical significance, biological relevance can be attached to very small changes in a marker. This is exemplified by reference to blood cholesterol levels (total cholesterol, LDL-cholesterol) in which, at the population level, a few percent change has large implications for the public health burden of coronary heart disease. The same applies to changes in blood pressure of only a few mm Hg [7]. Also, a minor gain in physical performance can have great effects in sport in which, at the top level, fractions of seconds may make the difference between success and failure [9]. Conversely, a change of several tens of percent in immune function parameters or stool weight, although perhaps significant statistically, may not have any biological relevance [13]. It is necessary that the conditions of both statistical significance and biological relevance are met if the outcome of a study is to provide support for a claim.

Criterion 6. A claim should be scientifically substantiated by taking into account the totality of the available data and by weighing of the evidence.

When assessing the validity of a claim, the reviewing bodies should have access to, and consider on their scientific merit, all relevant data.

The criteria are intended to ensure the scientific quality of studies and evidence to be used for the substantiation of claims. However, in many cases, results from individual studies may allow different interpretations or provide conflicting evidence. The quality of individual studies may differ and it is possible that not all research will be done to the highest, or even a common, standard. This can be due to the complexities of research in humans but also because data to support a claim may be used opportunistically from studies which had a different primary objective. There may however be a complementarity between individually incomplete studies which allows an assessment of the totality of the evidence to substantiate a claim. Conversely, a review of all studies taken together may reveal evidential inconsistencies that are not apparent from the review of a single study in isolation. The types of studies and evidence which can contribute to the substantiation of a claim are discussed under Criterion 2 and summarised in Table 2 (page I/13).

Selective presentation or consideration of studies and their outcomes is acceptable only if this is transparent and done on the basis of the quality of the data, for example if the selection of data is based on principles described in the commentary to these criteria.

In the evidence, overall, there should ideally be:

- consistency of results across the various categories of evidence and methodologies;
- valid dietary methods;
- randomised sampling;
- a dose response relationship between intakes of food or food components and the effects and health effect, if relevant;
- biological plausibility;

with all data supported by the use of valid markers (see criteria 3, 4 and 5).

Selective presentation of data depending on whether or not they would support the claim is not acceptable.

The evaluation of the available data may leave some questions unanswered. In such cases it should be considered whether these questions need to be answered by additional research, or whether or not the evidence overall supports the proposed claim.

All published studies should be reviewed and unpub-

lished data, including those that have been held back from publication for reasons of confidentiality, must also be considered.

Concluding comments and discussion

A set of criteria has been developed which defines the requirements for data submitted in the scientific substantiation of claims made on foods (see box below).

Criteria for the scientific substantiation of claims

1. The food or food component to which the claimed effect is attributed should be characterised.
2. Substantiation of a claim should be based on human data, primarily from intervention studies the design of which should include the following considerations:
   - 2(a) Study groups that are representative of the target group.
   - 2(b) Appropriate controls.
   - 2(c) An adequate duration of exposure and follow up to demonstrate the intended effect.
   - 2(d) Characterisation of the study groups' background diet and other relevant aspects of lifestyle.
   - 2(e) An amount of the food or food component consistent with its intended pattern of consumption.
   - 2(f) The influence of the food matrix and dietary context on the functional effect of the component.
   - 2(g) Monitoring of subjects' compliance concerning intake of food or food component under test.
   - 2 (h) The statistical power to test the hypothesis.
3. When the true endpoint of a claimed benefit cannot be measured directly, studies should use markers.
4. Markers should be:
   - biologically valid in that they have a known relationship to the final outcome and their variability within the target population is known; methodologically valid with respect to their analytical characteristics.
5. Within a study the target variable should change in a statistically significant way and the change should be biologically meaningful for the target group consistent with the claim to be supported.
6. A claim should be scientifically substantiated by taking into account the totality of the available data and by weighing of the evidence.

The criteria have been subjected to rigorous peer reviews by groups comprised of a broad-base of scientific and regulatory experts in three successive workshops. The criteria constitute a scientifically robust tool for evaluating the quality of data submitted in support of claims.

The PASSCLAIM Concerted Action has involved extensive collaboration and debate amongst different sectors including scientists and related expertise from academia and research institutes, industry, consumer interests and regulatory bodies. It has been elaborated by a process which has drawn on examples of existing best practice in respect of the use of investigative studies to monitor several health and well-being states and the reduction of disease risk, and of existing regulatory and advisory processes for the evaluation of claims.

The action has produced a consensus on the objective and transparent assessment of scientific evidence submitted to support a claim related to a food or food component. This approach is broken down to core issues that describe the context within which claims need to be considered, and into separate criteria that will facilitate the objective assessment and assist in the compilation of guidelines on the preparation of submissions. It emphasises that the overall consistency and coherence of all the evidence, i. e. the totality of the evidence, should be assessed. This approach should help those who are submitting evidence as well as those who are responsible for evaluating it, and this structure should also enable feedback to those submitting portfolios of evidence.

Thus this practical framework for the evaluation of scientific dossiers supporting claims can be expected to expedite and improve the efficiency of the regulatory review processes. It is hoped that this would give the European food manufacturing industry a competitive edge in the global market both from the establishment of claims, and also from an improved science base that this process might be expected to generate. This integrated strategy addresses consumer concerns and will assist in generating more consumer confidence in science-based claims on foods. Consumers should benefit through the availability of more foods with substantiated claims.

In the above respects the PASSCLAIM Concerted Action has met its objectives. Nonetheless, the action has identified other issues that need to be addressed. An important point that should be appreciated is that the template for the evaluative process, in its present form as it emerges from the PASSCLAIM process, essentially provides only guidance. The template needs to be applied intelligently and sensitively on a case by case basis with respect both to gaps in knowledge and to the development of new knowledge. It is to be expected that assessment of, for example, the validity of markers, study designs and the influence of dietary matrices on the effects of active components will require expert advice. Assessment of the totality, consistency and complementarity of evidence and the extrapolation of demonstrated benefits across gender and generation groups will also require expert judgement. Thus there will still

be a need for informed scientific advice in the regulatory process.

The systematic analyses of existing and potential claims carried out by expert groups during the course of the PASSCLAIM exercise have resulted in reviews of the availability of indicators of health and disease states within their respective areas of expertise [5–7, 9–12]. They have demonstrated the limitations of existing markers and have identified the need for better markers. In particular, the development of genomic and post-genomic molecular biology would be expected to improve the characterisation of populations, and the early detection of responses to interventions with foods and food components. The availability of such markers may facilitate the substantiation of claims by enabling more practicable and cost-efficient study protocols and timescales.

Nonetheless, the scientific substantiation of claims according to the PASSCLAIM criteria might require substantial and expensive studies in humans that would therefore, at a first glance, appear possible only for large companies who have the relevant economic and personnel resources. This may be particularly true for product specific claims but the criteria are also applicable to the substantiation of generic claims that can be made on a range of products containing the active food component.

PASSCLAIM agreed that the evidence required to support nutrient function, enhanced function, and reduction of disease risk claims needs to be of similar quality, and that as such these claims could be related to the schema developed in the previous concerted action on functional food science in Europe (FUFOSE). Nutrient function claims were not considered in FUFOSE but are now generally regarded as health claims (see Table 1). The particular issue relating to this spectrum of claims (see Fig. 4) is that they are in practice a continuum, and that it can be expected that on some occasions ambiguities and difficulties will arise in classifying claims that are submitted for approval. In essence Nutrient Function Claims will draw for substantiation on a broad "generally accepted base" such as that expressed recently in a WHO report [37], whereas Enhanced Function Claims will be more specific and will need "specific scientific studies" for their support. Disease Reduction Claims may need to draw on the broad spectrum of scientific data. However, there is no definite rule, each claim would need to be assessed in its own right.

There are some broader, more political, implications arising from this document.

Firstly, given the resource implications of developing and supporting enhanced function and disease risk reduction claims, it should be expected that producers will seek support to enable them to assert intellectual property rights for their innovations. As the regulatory environment for claims develops, this aspect will need to be considered, if the competitiveness of the EU food industry and the incentive for its investment in healthy foods are to be maximised.

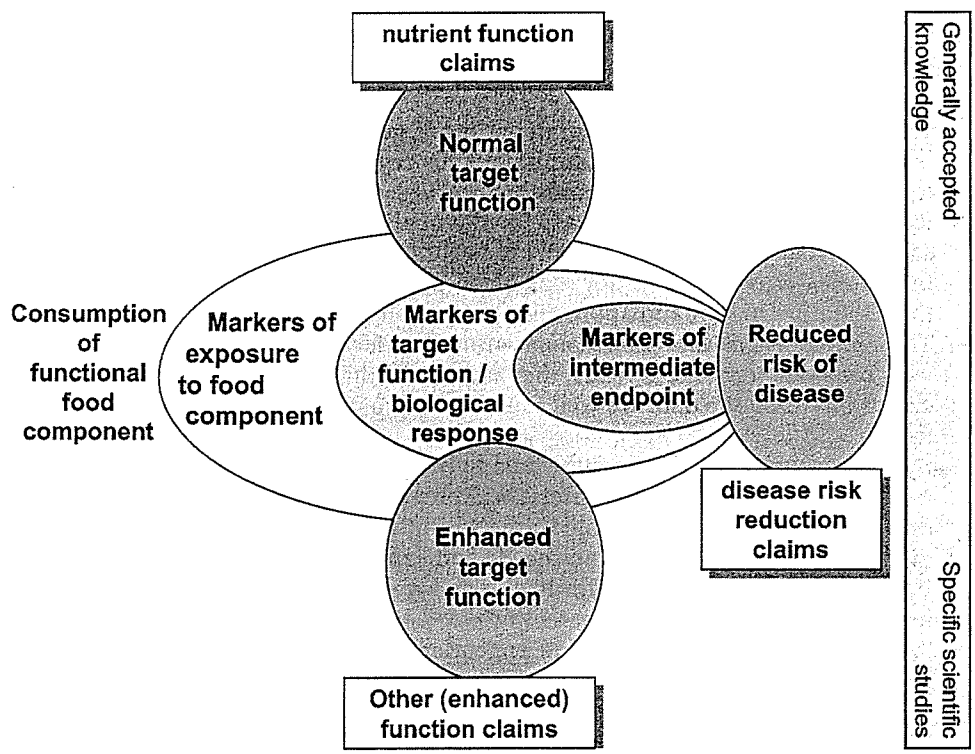

Fig. 4 Relationship between health claims addressed by PASSCLAIM and the FUFOSE concept of underlying scientific evidence

Secondly, it is appreciated that the criteria would be useful for innovative SMEs at an early stage of development of functional foods, in order to judge the feasibility of developing new products. There may be a need to identify common approaches to establishing the science bases for claims, which can be shared by large companies and SMEs to the benefit of both sectors. This may mean sharing resource and other means of collaboration. There may be a strategic need for competent authorities to support SMEs by investing in scientific support and networks, e. g. to undertake human nutrition studies.

Consumer confidence in claims is a key issue, from the producers' as well as from the consumers' points of view. Defining common criteria for the scientific substantiation of claims, supported by a broad group of European scientists representing both academia and industry, is an important step in establishing an environment in which consumers can be assured that claims made on foods are well-founded. Well-founded claims and associated explanations will contribute to consumer education. Consumer nutritional insight and knowledge will increase, and resultantly such informed consumers will be more able to choose products with benefits for health and well-being. In this way, claims substantiated in agreement with the PASSCLAIM criteria will contribute broadly to healthier diets for Europeans, and thereby to a decrease in the burden of diet-related diseases.

In summary, a number of potential benefits follow from these criteria. Achievement of these will require action to be taken to bring the criteria to a wider audience:

- The criteria provide a scientific framework that will facilitate the assessment of scientific support for claims on foods.
- This, in turn, will enable the compilation of guidelines on the preparation of submissions for regulatory review and approval of claims on foods.
- By establishing a robust standard for the quality of scientific data submitted in support of health claims, the criteria provide a basis for the harmonisation of the regulatory review and approval of such claims.
- The compliance of data submissions with the criteria will provide consumers with the assurance that claims based on the data are well founded and justified.
- By establishing a standard for the data to be submitted in support of claims, the criteria will provide the agri-food industry with a stable frame within which new products to meet consumer needs and expectations for foods with benefits for health and well-being can be developed.
- Systematic use of the criteria will engender a more informed use of scientific data in support of claims.

Acknowledgements ILSI Europe thanks all the collaborators in the PASSCLAIM Concerted Action (see list below) for providing their expertise, sharing their knowledge and devoting their time to produce this final PASSCLAIM publication. We thank the ILSI Europe Functional Food Task Force members and the PASSCLAIM Steering Committee members for their support and guidance. ILSI Europe is very grateful to Prof. Aggett, Prof. Asp, Dr. Antoine, Dr. Bellisle, Prof. German, Dr. Müller, Prof. Rechkemmer and Dr. Verhagen for their continuous input during more than four years, which led to a successful completion of the objectives of this Concerted Action. Finally, we thank the European Commission for positively evaluating this submission, providing continuous support to the various needs that arose during the course of the project and supporting financially the whole endeavour.

ILSI Europe

Coordinator:	Dr. Laura Contor
Scientific supervisor:	Dr. Loek Pijls
Project manager:	Ms. Sandra Tuijtelaars, M.Sc.
EC Responsible:	Dr. Jürgen Lucas, European Commission, DG Research, Unit E.2 Health, Food and Environment

The following people have contributed to the PASSCLAIM project:

Dr. Aline Adam (Cerestar Health and Nutrition Group, BE)
Mr. Tage Affertsholt (3A Business Consulting, DK)
Prof. Peter J. Aggett (University of Central Lancashire, GB)
Dr. Jean-Michel Antoine (Groupe Danone, FR)
Prof. Antti Aro (National Public Health Institute, FI)
Dr. Margaret Ashwell (Ashwell Associates, GB)
Prof. Nils-Georg Asp (SNF Swedish Nutrition Foundation, SE)
Prof. Michael Aviram (Rambam Medical Center and Technion Faculty of Medecine, IL)
Prof. Fernando Azpiroz (Hospital General Vall d'Hebron, ES)
Prof. Diána Bánáti (Central Food Research Institute – KEKI, HU)
Prof. Christian A. Barth (Verein für Nutrigenomik, DE)
Dr. Reto Battaglia (SQTS – Swiss Quality Testing Services, CH)
Dr. France Bellisle (Institut National de la Recherche Agronomique (INRA), FR)
Dr. Nino M. Binns (McNeil Consumer Nutritionals Europe, GB)
Dr. Richard Black (ILSI North America, US)
Dr. Hervé Blottière (Institut National de la Recherche Agronomique (INRA), FR)
Dr. Max Blum (DSM Nutritional Products, CH)
Dr. Stephanie Blum (Nestlé Research Centre, CH)
Prof. John E. Blundell (University of Leeds, GB)
Prof. Jean-Philippe Bonjour (Hôpitaux Universitaires de Genève, CH)
Ms. Antoinette Booyzen (ILSI South Africa, ZA)
Prof. Ingvar Bosaeus (Sahlgrenska University Hospital, SE)
Dr. Marlou Bosch (Cosun Food Technology Centre, NL)
Dr. Raphaelle Bourdet-Sicard (Danone Vitapole, FR)
Dr. Julio Boza Puerta (Puleva Biotech S. A., ES)
Dr. Francesco Branca (Italian National Institute for Research on Food and Nutrition, IT)
Prof. Per Brandtzaeg (University of Oslo, NO)
Prof. Furio Brighenti (University of Parma, IT)
Dr. Fred Brouns (Cerestar R& D Centre, BE)
Prof. Åke Bruce (National Food Administration, SE)
Prof. Philip Calder (University of Southampton, GB)
Prof. Nancy M. Childs (Saint Joseph's University, US)
Ms. Merete Myrup Christensen (Arla Foods Innovation, DK)
Dr. Roberto Ciati (Barilla G&R F.lli Spa, IT)
Dr. Andrew Cockburn (Monsanto UK Ltd., UK)
Mr. Pierre Combris (Institut National de la Recherche Agronomique (INRA), FR)
Dr. Laura Contor (ILSI Europe, BE)
Prof. Cyrus Cooper (MRC – Southampton General Hospital, GB)

Prof. John H. Cummings (University of Dundee, GB)
Dr. Karen Cunningham (Coca-Cola Great Britain & Ireland, GB)
Prof. Dr. Hannelore Daniel (Technical University of Munich, DE)
Ms. Ellen de Ree (Numico Research, NL)
Dr. Jan de Vries (Royal Friesland Foods, NL)
Prof. Pierre Déchelotte (CHU de Rouen, FR)
Prof. Nathalie Delzenne (Université Catholique de Louvain, BE)
Mr. Gontran Dondain (Colloïdes Naturels International, FR)
Dr. Laurence Doughan (SPF Santé publique, Sécurité de la Chaîne Alimentaire et Env, BE)
Ms. Selma Doyran (FAO – Food and Agriculture Organization, IT)
Prof. Lars Ove Dragsted (Danish Institute for Food and Veterinary Research, DK)
Mr. Nigel Duffin (So Good International, GB)
Mr. Ruedi Duss (CreaNutrition, CH)
Dr. Damian Edwards (Masterfoods, GB)
Prof. Dr. Aziz Eksi (Ankara University, TR)
Dr. Muriel Eliaszewicz (Agence Française de Sécurite Sanitaire des Aliments, AFSSA, FR)
Prof. Dr. Ibrahim Elmadfa (University of Vienna, AT)
Mr. Reg J. Fletcher (Kellogg's Company of Great Britain Ltd., GB)
Prof. Albert Flynn (University College Cork, IE)
Dr. Mikael Fogelholm (The UKK Institute, FI)
Dr. Anne Franck (Raffinerie Tirlemontoise – Orafti, BE)
Prof. Keith N. Frayn (University of Oxford, GB)
Mr. Fumitake Fukutomi (ILSI Japan, JP)
Dr. Michèle Garabedian (Hôpital Saint-Vincent de Paul, FR)
Prof. Bruce German (Nestlé Research Centre, CH)
Prof. Michael J. Gibney (Institute of European Food Studies, IE)
Prof. Glenn R. Gibson (University of Reading, GB)
Prof. Michael Gleeson (Loughborough University, GB)
Dr. Mirjam Govers (Numico Research, NL)
Ms. Christine Gratus (SACN, GB)
Dr. Bruce A. Griffin (University of Surrey, GB)
Prof. Rolf Grossklaus (Federal Institute for Risk Assessment (BfR), DE)
Dr. Christian Grugel (Bundesanstalt für Gesundheitlichen Verbraucherschutz, DE)
Dr. Francisco Guarner (Hospital General Vall d'Hebron, ES)
Mr. Nicolas Guggenbuhl (Diffu-Sciences Sprl., BE)
Dr. Bernd Haber (Nutrinova GmbH, DE)
Dr. Thomas Hatzold (Kraft Foods R&D, Inc., DE)
Prof. Joseph Hautvast (Health Council of the Netherlands, NL)
Dr. Henk. F. J. Hendriks (TNO – Nutrition and Food Research Institute, NL)
Dr. Peter Hespel (Katholic University Leuven, BE)
Ms. Rosemary Hignett (Food Standards Agency, GB)
Dr. Peter Hollman (RIKILT Institute of Food Safety, NL)
Mr. John Howlett (The Kapajo Group, GB)
Ms. Sally-Ann Hughes (Glaxo Smith Kline, GB)
Prof. Erika Isolauri (University of Turku, FI)
Ms. Anna Jacob (ILSI Southeast Asia, SG)
Dr. Asker Jeukendrup (University of Birmingham, GB)
Dr. Ian Johnson (Institute of Food Research, GB)
Prof. Hans-Georg Joost (DIFE – German Institute of Human Nutrition, DE)
Ms. Anna Jung (European Food Information Council (EUFIC), BE)
Prof. Anthony Kafatos (University of Crete – School of Medicine, GR)
Prof. Dr. K. Wolfgang Kallus (Karl-Franzens-Universität Graz, AT)
Dr. Jaap Keijer (RIKILT Institute of Food Safety, NL)
Dr. Michele B. Kellerhals (Beverages Partners Worldwide, CH)
Ms. Beate Kettlitz (European Consumers Organisation (BEUC), BE)
Ir. Jaap D. Kluifhooft (DSM Food Specialties, NL)
Dr. Michael E. Knowles (Coca-Cola, BE)
Prof. Dr. Frans J. Kok (Wageningen University, NL)
Dr. Riitta Korpela (Valio Ltd., FI)
Prof. Robert Kroes (Utrecht University – IRAS, NL)
Dr. Miltos Ladikas (Lancaster University, GB)
Dr. Liisa Lähteenmäki (VTT Biotechnology, FI)
Dr. Janet Lambert (Coca-Cola, GB)
Ms. Aude L'Hirondel (Euro Coop, BE)
Dr. Barbara Livingstone (University of Ulster, GB)
Dr. Jürgen Lucas (European Commission, BE)
Ms. Carina Madsen (ILSI Europe, BE)
Dr. Giancarlo Malgarini (Ferrero Group, BE)
Dr. Ascensión Marcos Sánchez (Consejo Superior de Investigaciones Cientificas (CSIC), ES)
Dr. Paule Martel (Institut National de la Recherche Agronomique – I. N. R. A., FR)
Prof. Ron Maughan (Loughborough University, GB)
Dr. Annika Mäyra-Mäkinen (Raisio Group Plc, FI)
Prof. Dr. Ronald P. Mensink (Maastricht University, NL)
Dr. P. Diederick Meyer (Sensus, NL)
Ms. Angelika Michel-Drees (Federation of German Consumer Organisations, DE)
Dr. Hubert Milon (Nestlé Research Centre, CH)
Ms. Daniela Muchna (Red Bull GmbH, AT)
Dr. Detlef J. G. Müller (Procter & Gamble Service GmbH, DE)
Dr. Marja Mutanen (University of Helsinki, FI)
Ms. Sabine Nafziger (Confederation of the Food and Drink Industries of the EU (CIAA), BE)
Mr. Guy Nantel (Food and Agriculture Organization (FAO), IT)
Dr. Ute Obermüller-Jevic (BASF Aktiengesellschaft, DE)
Dr. John O'Brien (Food Safety Authority of Ireland, IE)
Dr. Daphne Pannemans (Former ILSI Europe, BE)
Prof. Jana Parizkova (Institute of Endocrinology, CZ)
Dr. Gérard Pascal (Institut National de la Recherche Agronomique (INRA), FR)
Dr. Christoph Persin (Kampffmeyer Mühlen GmbH, DE)
Dr. Loek Pijls (ILSI Europe, BE)
Dr. Rudolf Poledne (Institute of Clinical & Experimental Medicine, CZ)
Prof. Dr. Beatrice Pool-Zobel (Friedrich-Schiller University of Jena, DE)
Dr. Ann Prentice (MRC – Human Nutrition Research, GB)
Prof. Joseph J. Rafter (Karolinska Institute, SE)
Prof. Dr. Gerhard Rechkemmer (Technische Universität München, DE)
Prof. Gabriele Riccardi (Università Federico II – Napoli, IT)
Prof. David P. Richardson (DPR Nutrition Limited, GB)
Dr. Blanka Richterova (Charles University, CZ)
Ms. Catherine Rioux (DGCCRF, FR)
Prof. Marcel Roberfroid (BE)
Dr. Pilar Rodríguez Iglesias* (European Food Safety Authority, BE)
Dr. Peter J. Rogers (University of Bristol, GB)
Dr. Gun Roos (SIFO – National Institute for Consumer Research, NO)
Dr. Richard Ross (Glaxo SmithKline, GB)
Prof. Ian Rowland (University of Ulster, GB)
Ms. Melanie Ruffell (Joint Health Claims Initiative, GB)
Dr. Jiri Ruprich (National Institute of Public Health, CZ)
Prof. Seppo Salminen (University of Turku, FI)
Prof. Wim H. M. Saris (Maastricht University, NL)
Mr. Matthias Sass (Rudolf Wild GmbH & Co. KG, DE)
Prof. Dr. Ir. Gertjan Schaafsma (TNO – Nutrition and Food Research Institute, NL)
Ms. Peggy Schuhmann (Swiss Quality Testing Services (SQTS), CH)
Dr. Mauro Serafini (Italian National Institute for Research on Food and Nutrition, IT)
Dr. Toshio Shimizu (Fresco Japan, JP)
Prof. Colette Shortt (Yakult UK Ltd., GB)
Prof. Wilhelm Stahl (Heinrich-Heine-Universität Düsseldorf, DE)
Prof. Dr. Peter Stehle (Friedrich-Wilhelms-University Bonn, DE)
Dr. Jan M. Steijns (Campina, NL)
Dr. Vladimír Stich (Charles University, CZ)
Dr. Julian Stowell (Danisco Sweeteners, GB)
Mr. Kazuo Sueki (ILSI Japan, JP)

* Observer.

Prof. Linda Tapsell (University of Wollongong, AU)
Dr. Christine Taylor (World Health Organization (WHO), CH)
Dr. E. Siong Tee (Nutrition Society of Malaysia, MY)
Dr. Hans-Ulrich Ter Meer (Nutrinova, DE)
Dr. Stephan Theis (Südzucker AG, DE)
Dr. Kraisid Tontisirin (Food and Agriculture Organization (FAO), IT)
Ms. Sandra Tuijtelaars (ILSI Europe, BE)
Prof. Hely Tuorila (University of Helsinki, FI)
Dr. Dominique Turck (Hôpital Jeanne de Flandre, FR)
Dr. Bert Uijtewaal (Bayer CropScience, BE)
Ms. Louise Ungerth (Konsumentföreningen Stockholm, SE)
Dr. Nico van Belzen (ILSI Europe, BE)
Dr. Henk van den Berg (Stichting Voedingscentrum Nederland, NL)
Mr. Pier van der Kouwe (Bayer CropScience, BE)
Dr. Annémiek van der Laan (Consumentenbond, NL)
Prof. Dr. Henk van Loveren (RIVM – National Institute of Public Health & the Environment, NL)
Dr. Louis van Nieuwland (Voedingscentrum, NL)
Dr. Jan Van Loo (Raffinerie Tirlemontoise – Orafti, BE)
Prof. Hans van Trijp (Unilever, NL)
Ms. Marian Verbruggen (Acatris Holding BV, NL)
Dr. Hans Verhagen (Unilever Health Institute, NL)
Drs. Paulus M. Verschuren (Unilever Health Institute, NL)
Prof. Bengt Vessby (University of Uppsala, SE)
Mr. Volker Viechtbauer (Red Bull, AT)
Prof. Dr. Burckhard Viell (Federal Institute for Risk Assessment (BfR), DE)
Dr. Bernhard Watzl (Federal Research Centre for Nutrition and Food, DE)
Dr. Peter Weber (DSM Nutritional Products Ltd, CH)
Dr. Hans-Ulrich Wekel (BASF, DE)
Prof. Dr. Joachim Westenhoefer (Hamburg University of Applied Sciences, DE)
Dr. Ingmar Wester (Raisio Benecol Ltd, FI)
Dr. Peter Williams (University of Wollongong, AU)
Ms. Boon Yee Yeong (ILSI Southeast Asia, SG)
Dr. Elizabeth A. Yetley (The National Institutes of Health, US)
Ms. Christina Zehaluk (Health Canada, CA)

Glossary

Bioavailability: The fractional amount of a nutrient or other bioactive substance that, after ingestion, becomes available for use in target tissues.

Case-control study: Study that compares the exposure to a suspected cause of a disease in people with that disease (the cases) to the exposure in those without that disease (controls); exposure is thus assessed retrospectively. See also 'cross-sectional study'.

Claim: Any message or representation, including pictorial, graphic or symbolic representation, which states, suggests or implies that a food has particular characteristics.

Clinical study: Study of any therapeutic, prophylactic, or diagnostic agent conducted in human subjects, in which the clinical and statistical description, presentations, and analysis are fully integrated into a single report.

Codex Alimentarius: Literally: 'Food Code'. An organisation that creates and compiles standards, codes of practice and recommendations. Membership is open to all countries associated with the Food and Agricultural Organization of the United Nations and with the World Health Organization. At present (2005) Codex has 168 members and covers more than 98% of the world's countries. Also non-governmental organisations have input in Codex (www.codexalimentarius.net).

Cohort study: Prospective observational study in which data on exposure to suspected causes of e.g. a disease are collected in a selected/recruited group of people who do not yet have the disease(s) under investigation. The subjects are then followed for a period of time, after which it can be assessed whether development of disease is related to the (presence of) suspected causes.

Confounding factors, confounders: A certain exposure may be associated with a disease or other outcome, without this association being causal. This can result from a third factor being a cause of both; such a factor is referred to as 'confounder'. In other words: an alternative cause for the disease in question that is unequally distributed among those exposed and non-exposed to the putative agent (Hayes 2001 in FOSIE [3]).

Cross-sectional study: A study design that relates the rates of a certain exposure to the levels of an outcome of interest in a number of individuals or populations. Key feature is that exposure and outcome are measured at the same point in time.

Disease risk reduction claim: A claim that states or implies that consumption of a product reduces the risk of occurrence of a certain disease. See also 'enhanced function claim', 'health claim', 'medical claim' and 'prevention of disease'.

Dose-response relationship: The finding that the level of variable A changes as changes in the level of variable B occur. 'A' may be the level of a function or parameter in the body, or the risk of a disease and 'B' may be the intake of a food component. The existence of such a relationship adds to the probability that the observed relationship is causal.

Endpoint: A variable or outcome that is relevant in itself, e.g. survival time after medical surgery, time to run a marathon, fewer periods of gastrointestinal discomfort, or a reduced risk of a disease. The level of a surrogate or intermediate endpoint – also referred to as 'marker' – is in itself not relevant, but is indirectly relevant because it reflects a relevant endpoint. See also 'marker'.

Enhanced function claim: A claim that states or implies that the consumption of a product enhances a bodily function. 'Enhanced' aims to distinguish effects on functions other than the currently well-established effects of nutrients (so-called 'nutrient function claims'). As a result, a newly discovered effect on a function may initially give rise to an 'enhanced function claim', whereas once well established it would render a 'nutrient function claim'. See also 'disease risk reduction claim', 'health claim' and 'medical claim'.

Epidemiology: The study of health and the occurrence of diseases and their predictors and causes.

Food: Material used in the body to sustain growth, repair and other vital processes [38]. That which can be eaten... to stay alive and to grow ([39]). Any substance or product, ..., intended to be ingested by humans. 'Food' includes drink, ... [40].

Food component: components such as ingredients and food additives intentionally added to foods, and also components inherently present as part of the essential composition of foods.

FUFOSE: "Functional Food Science in Europe"; a European Commission Concerted Action, coordinated by ILSI Europe and completed in 1999 [2].

Generic claim: A claim based on knowledge from evidence generally available in the scientific literature and/or on recommendations from national or international public health bodies.

Glycaemic index: The glycaemic index is defined as the incremental

area under the blood glucose response curve of a 50 g carbohydrate portion of a test food expressed as a percent of the response to the same amount of carbohydrate from a standard food taken by the same subject [41].

Good clinical practice (GCP): a standard for the design, conduct, performance, monitoring, auditing, recording, analyses, and reporting of clinical trials that provides assurance that the data and reported results are credible and accurate, and that the rights, integrity, and confidentiality of trial subjects are protected [42].

Good laboratory practice (GLP): a set of principles that provides a framework within which laboratory studies are planned, performed, monitored, recorded, reported and archived. These studies are undertaken to generate data by which the hazards and risks to users, consumers and third parties, including the environment, can be assessed for pharmaceuticals, agrochemicals, cosmetics, food and feed additives and contaminants, novel foods and biocides. GLP helps assure regulatory authorities that the data submitted are a true reflection of the results obtained during the study and can therefore be relied upon when making risk/safety assessments [43].

Health: a state of complete physical, mental and social well-being and not merely the absence of diseases or infirmity [44].

Health claim: Any representation that states, suggests, or implies that a relationship exists between a – constituent of a – food and health (www.codexalimentarius.net/reports.asp). See also 'disease risk reduction claim', 'enhanced function claim' and 'medical claim'.

Intervention study: Study in which investigators intervene by allocating and establishing one or more treatments ("interventions") to or in certain subjects. See also 'observational study'. See also 'randomised controlled trial'.

Marker: A variable that is of interest because it marks or reflects a certain phenomenon of interest. One preferably avoids the confusing terms 'surrogate marker' and 'intermediate marker'. See also 'endpoint' and 'valid'.

To match: To be equal to; corresponding with regard to certain characteristics [39]. A method used to create study groups that are maximally similar, in order to ascribe differences in outcome to a certain factor in which the groups do differ. In e. g. a case control study one may 'match' controls to the identified cases by selecting a group of other patients in the cases' hospital who do not have the disease under study, but have similar age, ethnic background and gender. See also 'randomise'.

Matrix: Substance in which something is embedded [39].

Medical/medicinal claim: A claim (see 'claim') that states or implies that a food or a food component has the property of treating, preventing or curing human disease or makes any reference to such property. 'Human disease' means any injury, ailment or adverse condition, whether body or mind. Such claims are prohibited on foods; this prohibition creates the legal separation between foods and medicines. See also 'disease risk reduction claim', 'enhanced function claim', 'health claim' and 'prevention of disease'.

Meta-analysis: A quantitative summary of several individual studies of a similar type. Both intervention and observational studies can be meta-analysed. See also 'pooled analysis'.

Nocebo: see 'Placebo'.

Nutrient function claim: A claim that describes the physiological role of a nutrient in growth, development and normal functions of the body.

Nutrition: The act or process of nourishing; the process by which foods are taken in and utilised by the body for growth, normal function and maintenance of health [38].

Observational: From 'to observe': to see and notice; to watch carefully [39]. In an observational study, researchers do not intervene but only observe outcomes of interest and – the levels of – their suspected causes, e. g. cohort or case-control study. See also 'cross-sectional study' and 'intervention study'. Observational studies are often loosely referred to as epidemiological studies.

Placebo: an inert or innocuous substance used especially in controlled experiments testing the efficacy of another substance (as a drug) [38]. A "placebo" is especially useful to control for any beneficial effect that would occur in an experiment (due to the testing conditions themselves) but that would not be caused by the active agent in the tested food or food ingredient. Alternatively, a "nocebo" effect (an undesirable consequence induced by the particular test conditions) can also occur and should be discriminated from the action of the active substance under test.

Pooled analysis: An analysis of the combined, original data of several individual studies. See also 'meta-analysis'.

(Statistical) Power: The minimum size effect that can be demonstrated with statistical significance, given a certain study design and sample size. Based on the power required, one *a priori* calculates the sample size, and hence the study size, needed to achieve that. See also 'statistical significance'.

Prevention of disease: Hindrance [39] of the onset of disease. This hindrance may reduce the probability or risk of a disease to zero, but in diet-related diseases it usually reduces the risk to a lesser degree. See also 'disease risk reduction claim' and 'medical claim'.

P(robability)-value: The probability of observing in a subgroup or sample – by chance – an effect (a difference, an association) of minimally a certain size, in the situation that the effect does actually not exist in the original or overall population. See also 'statistical significance'.

Product-specific claim: A claim that a relationship exists between a specific food product, or a component of a specific food product, and health.

Randomisation: In intervention studies subjects may be randomly (i. e. determined by fate/chance) allocated either to undergo a certain intervention or to be part of a control group (or to undergo another intervention). Purpose of randomisation is to create groups that are likely to differ only with regard to the intervention under study. As a result, the effects observed can principally be ascribed to the intervention. See also 'to match'.

Randomised controlled trial (RCT): Study design in which subjects are randomly allocated to study groups. As a result the groups will expectedly not differ systematically, except with regard to an intervention that one group will undergo and the other will not. As a result, the effects observed can principally be ascribed to the intervention. See also 'intervention study' and 'to randomise'.

Representative: Serving as an example of a class or group; typical specimen of a group [39]. A sample out of a larger group is representative in certain aspects for that larger group if it does not differ systematically from that group in these aspects; if it is typical for that group.

Risk: Probability or chance of meeting a certain – usually unwanted – event [39]. The probability of loss or peril [38]. The probability and severity of an adverse effect/event occurring to man or the environ-

ment following exposure, under defined conditions, to a risk source [40].

Statistical significance: If the p-value for a certain observed effect is smaller than, e. g., 5 %, the assumption or hypothesis that the effect does not exist is refuted. The observed effect is then referred to as 'statistically significant'. See 'p(robability)-value'. A statement about statistical significance is a generalisation of a probability from a sample to the universe from which it has been drawn [16].

Target function/variable: A bodily function that is a target for intervention and measurement, in the scope of maintainance or improvement of health, or reduction of risk of disease.

Well-being: A positive and sustainable state that allows individuals, groups or nations to thrive and flourish. At the level of an individual, well-being refers to psychological, physical and social states that are distinctively positive [45].

Annex 1: Sensitivity and specificity

In evaluating and selecting markers, the sensitivity and specificity of the marker are important. In studies with humans, sensitivity is commonly defined as the proportion of a population with a certain characteristic (e. g. disease, health status) that is correctly classified on the basis of measurements as subjects with that characteristic. In the following Table, sensitivity can be quantified as A/(A + C). A high sensitivity implies a low proportion of false-negatives (category C). A study is only successful, however, if the proportion of false-positives (category B) is small as well. Thus the study has to be specific as well, i. e. a large proportion of subjects without disease or health status are correctly classified as such: D/(B + D) must be high.

		Reality	
		+	-
test	+	A	B
	-	C	D

A number of true positives, *B* number of false positives, *C* number of false negatives, *D* number of true negatives, *Sensitivity* probability of a positive test in people with the disease (A/A + C), *Specificity* probability of a negative test in people without the disease (D/B + D), *Positive predictive value* probability of a person having the disease when the test is positive (A/A + B), *Negative predictive value* probability of a person not having the disease when the test is negative (D/C + D)

Annex 2: Accuracy, precision, repeatability, reproducibility, linear and dynamic range – as used in criterion 4

It is important to have an insight into the practical performance of an analytical method. Control of the analytical performance of measurements is a prerequisite for a good and true study result. A high repeatability and reproducibility can reduce the number of measurements that need to be done. Precision refers to how close measurements of the same quantity are to each other, even if they are not close to the true value. A high precision, in addition to knowledge on biological variation, can reduce the number of subjects needed in a study. Accuracy refers to how close a measurement is to the true value of what is being measured. A high accuracy allows comparison of data across laboratories. The linear and dynamic range determine how many data/subjects can be considered in the overall assessment of the results.

The distinction between accuracy and precision is illustrated in Fig. 5 in which the symbols distributed over the targets represent a series of measurements. A symbol positioned at the bull's eye represents a perfect measurement – a measurement giving a value exactly the same as the true value.

The official definitions of accuracy, precision, repeatability, reproducibility, linear range and dynamic range are according to IUPAC Compendium of Chemical Terminology 2nd Edition (1997); (http://www.iupac.org/publications/compendium/R.html). An additional reference for these quality criteria can be found in ISO norm 5725: (http://www.iso.ch/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER = 11837).

Accuracy (of measurement)

Closeness of the agreement between the result of a measurement and a true value of the measurand.

Notes:

1. Accuracy is a qualitative concept.
2. The term precision should not be used for accuracy.

Precision

The closeness of agreement between independent test results obtained by applying the experimental procedure under stipulated conditions. The smaller the random part of the experimental errors

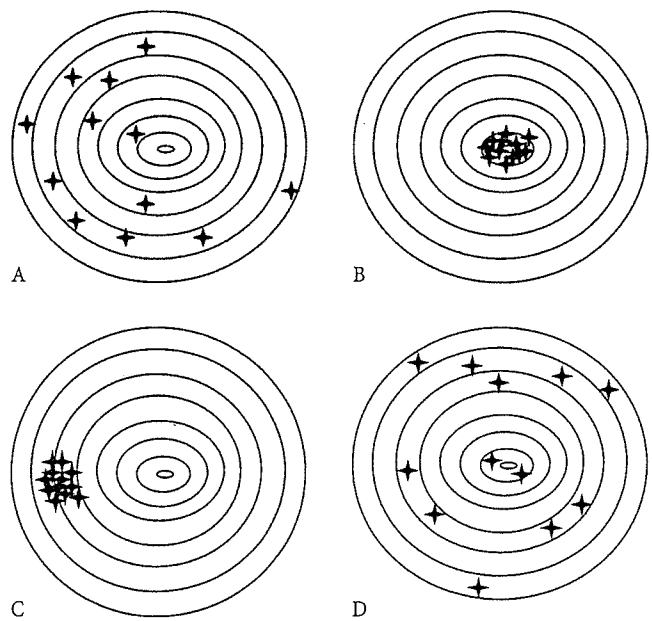

Fig. 5 Precision and accuracy. **A** Neither precise nor accurate. Since none of the darts are close to the bull's eye, the measurements they represent are not very accurate. Also, since the darts are not very close to each other, the set of five measurements here is not very precise either. **B** Both precise and accurate. The measurements are all close to the true value, so they are accurate. Also, the measurements are all close to each other, so they are precise. **C** Precise but not accurate. Since all of the measurements are close together, they are precise, but since they are not close to the true value, they are not accurate. **D** Accurate but not precise. The mean of all of the measurements is close to the true value, but since they are not very close together, they are not precise.

which affect the results, the more precise the procedure. A measure of precision (or imprecision) is the standard deviation.

Comment: Precision is sometimes misused for accuracy. This problem will be avoided if one recognizes that precision relates only to dispersion, not to deviation from the (conventional) true value. Imprecision has been defined as 'the standard error of the reported value.'

Repeatability

The closeness of agreement between independent results obtained with the same method on identical test material, under the same conditions (same operator, same apparatus, same laboratory and after short intervals of time). The measure of repeatability is the standard deviation qualified with the term: 'repeatability' as repeatability standard deviation. In some contexts repeatability may be defined as the value below which the absolute difference between two single test results obtained under the above conditions, may be expected to lie with a specified probability.

Reproducibility

The closeness of agreement between independent results obtained with the same method on identical test material but under different conditions (different operators, different apparatus, different laboratories and/or after different intervals of time). The measure of reproducibility is the standard deviation qualified with the term 'reproducibility' as reproducibility standard deviation.

In some contexts reproducibility may be defined as the value below which the absolute difference between two single test results on identical material obtained under the above conditions, may be expected to lie with a specified probability. Note that a complete statement of reproducibility requires specification of the experimental conditions which differ.

Linear range

Concentration range over which the intensity of the signal obtained is directly proportional to the concentration of the species producing the signal.

Dynamic range (of an analyser)

The ratio between the maximum usable indication and the minimum usable indication (detection limit). A distinction may be made between the linear dynamic range, where the response is directly proportional to concentration, and the dynamic range where the response may be non-linear, especially at higher concentrations.

References

1. Richardson DP, Affertsholt T, Asp N-G, Bruce A, Grossklaus R, Howlett J, Pannemans D, Ross R, Verhagen H, Viechtbauer V (2003) PASSCLAIM – Synthesis and review of existing processes. Eur J Nutr 42(Suppl 1):96–111
2. Diplock AT, Aggett PJ, Ashwell M, Bornet F, Fern EB, Roberfroid MB (1999) Scientific Concepts of Functional Foods in Europe: Consensus Document. Br J Nutr 81:S1–S27
3. Renwick AG, Barlow SM, Hertz-Picciotto I, Boobis AR, Dybing E, Edler L, Eisenbrand G, Greig JB, Kleiner J, Lambe J, Müller D, Smith M, Tritscher A, Tuijtelaars S, van den Brandt PA, Walker R, Kroes R (2003) Risk Characterisation of Chemicals in Food and Diet – Final part of Food Safety in Europe (FOSIE): Risk Assessment of Chemicals in Food and Diet. Food Chem Toxicol 41:1211–1271
4. Commission of the European Communities. Proposal for a Regulation of the European Parliament and of the Council on nutrition and health claims made on foods. COM(2003)424 final, 1–34. 2003. Luxembourg, Commission of the European Communities
5. Asp NG, Cummings JH, Mensink RP, Prentice A, Richardson DP (2003) Process for the Assessment of Scientific Support for Claims on Foods (PASSCLAIM) – Phase One: Preparing the Way. Eur J Nutr 42(Suppl 1):1–119
6. Asp NG, Cummings JH, Howlett J, Rafter J, Riccardi G, Westenhoefer J (2004) Process for the Assessment of Scientific Support for Claims on Foods (PASSCLAIM) – Phase Two: moving forward. Eur J Nutr 43(Suppl 2):1–183
7. Mensink RP, Aro A, Den Hond E, German JB, Griffin BA, Ter Meer H-U, Mutanen M, Pannemans D, Stahl W (2003) PASSCLAIM – Diet-related cardiovascular disease. Eur J Nutr 42(Suppl 1): 6–27
8. Prentice A, Bonjour J-P, Branca F, Cooper C, Flynn A, Garabedian M, Müller D, Pannemans D, Weber P (2003) PASSCLAIM – Bone health and osteoporosis. Eur J Nutr 42(Suppl 1):28–49
9. Saris WHM, Antoine J-M, Brouns F, Fogelholm M, Gleeson M, Hespel P, Jeukendrup AE, Maughan RJ, Pannemans D, Stich V (2003) PASSCLAIM – Physical performance and fitness. Eur J Nutr 42 (Suppl 1):50–95
10. Riccardi G, Aggett PJ, Brighenti F, Delzenne N, Frayn K, Niewenhuizen A, Pannemans D, Theis S, Tuijtelaars S, Vessby B (2004) PASSCLAIM – Body weight regulation, insulin sensitivity, and diabetes risk. Eur J Nutr 43(Suppl 2):7–46
11. Rafter J, Govers M, Martel P, Pannemans D, Pool-Zobel B, Rechkemmer G, Rowland I, Tuijtelaars S, Van Loo J (2004) PASSCLAIM – Diet-related cancer. Eur J Nutr 43(Suppl 2):47–84
12. Westenhoefer J, Bellisle F, Blundell JE, de Vries J, Edwards D, Kallus W, Milon H, Pannemans D, Tuijtelaars S, Tuorila H (2004) PASSCLAIM – Mental state and performance. Eur J Nutr 43(Suppl 2):85–117
13. Cummings JH, Antoine J-M, Azpiroz F, Bourdet-Sicard R, Brandtzaeg P, Calder PC, Gibson GR, Guarner F, Isolauri E, Pannemans D, Shortt C, Tuijtelaars S, Watzl B (2004) PASSCLAIM – Gut health and immunity. Eur J Nutr 43 (Suppl 2):118–173
14. Bellisle F, Diplock AT, Hornstra G, Koletzko B, Roberfroid M, Salminen S, Saris WHM (1998) Functional Food Science in Europe. Br J Nutr 80(Suppl 1):1–193
15. Hill AB (1965) The environment and disease: association or causation? Proc R Soc Med 58:295–300
16. Susser M (1986) The logic of Sir Karl Popper and the practice of epidemiology. Am J Epidemiol 124:711–718
17. Willett WC (1990) Nutritional epidemiology. Oxford University Press, New York
18. Bingham SA (1991) Limitations of the various methods for collecting dietary-intake data. Ann Nutr Metab 35: 117–127

19. Bingham SA, Gill C, Welch A, Kay K, Cassidy A, Khaw KT, Sneyd MJ, Key TJA, Roe L, Day NE (1994) Comparison of dietary assessment methods in nutritional epidemiology – weighed records V 24-h recalls, food-frequency questionnaires and estimated-diet records. Br J Nutr 72:619–643
20. Meyer, B (2002) Methods of Dietary Intake Assessment.www.fhcrc.org/phs/cvdeab/chpt17. html (17)
21. FAO/WHO (1996) Preparation and use of food-based dietary guidelines. Geneva, Food and Agriculture Organisation, World Health Organization
22. Bingham SA, Cassidy A, Cole TJ, Welch A, Runswick SA, Black AE, Thurnham D, Bates C, Khaw KT, Key TJA, Day NE (1995) Validation of weighed records and other methods of dietary assessment using the 24-h urine nitrogen technique and other biological markers. Br J Nutr 73:531–550
23. Bingham SA, Gill C, Welch A, Cassidy A, Runswick SA, Oakes S, Lubin R, Thurnham DI, Key TJA, Roe L, Khaw KT, Day NE (1997) Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers. Int J Epidemiology 26: S137–S151
24. Black AE, Welch AA, Bingham SA (2000) Validation of dietary intakes measured by diet history against 24 h urinary nitrogen excretion and energy expenditure measured by the doubly-labelled water method in middle-aged women. Br J Nutr 83:341–354
25. McKeown NM, Day NE, Welch AA, Runswick SA, Luben RN, Mulligan AA, McTaggart A, Bingham SA (2001) Use of biological markers to validate self-reported dietary intake in a random sample of the European Prospective Investigation into Cancer United Kingdom Norfolk cohort. Am J Clin Nutr 74: 188–196
26. Steptoe A, Perkins-Lorras L, McKay C, Rink R, Hilton S, Cappuccio FP (2003) Behavioural counselling to increase consumption of fruit and vegetable in low income adults: randomised trial. Br Med J 326:855
27. Margetts B, Nelson M (1997) Design Concepts in Nutritional Epidemiology, 2nd Edition. Oxford University Press, Oxford
28. Cameron ME, Van Staveren WA (1988) Manual on methodology for food consumption studies. Oxford University Press, New York
29. van Staveren W, Ocke M (2001) Estimation of dietary intake. Present knowledge of Nutrition. In: Bowman BA, Russell RM (eds) 8th edn. ILSI Press, Washington, DC, pp 605–616
30. Tooze JA, Subar AF, Thompson FE, Troiano R, Schatzkin A, Kipnis V (2004) Psychosocial predictors of energy underreporting in a large doubly labeled water study. Am J Clin Nutr 79:795–804
31. Moher D, Schultz KF, Altmann DG, Consort Group (Consolidated Standards of Reporting Trials) (2001) The CONSORT statement: revised recommendation for improving the quality of reports of parallel-group randomised trials. Lancet 357:1191–1194
32. Truswell AS (2001) Levels and kinds of evidence for public-health nutrition. Lancet 357:1061–1062
33. Commission of the European Communities. Directive 2001/20/EC on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. Official Journal of the European Communities. L121, 34–44. 2001. Luxembourg, Commission of the European Communities
34. Hulka BS, Wilkosky TC, Griffith JD (1990) Biological markers in epidemiology. Oxford University Press, New York
35. van Poppel G, Verhagen H, Heinzow B (1997) Biomarkers in epidemiological and toxicological nutrition research. In: Tennant D (ed) Food Chemical Risk Analysis. Chapman and Hall, London, pp 87–108
36. Health Council of the Netherlands (2003) Foods and dietary supplements with health claims. Publication nr 2003/09E. The Hague, Health Council of the Netherlands
37. World Health Organization (2003) Diet, nutrition and the prevention of chronic diseases. Report of a Joint WHO/FAO Expert Consultation Geneva, 28 January – 1 February 2002. 916, 1–160. Geneva, World Health Organization. WHO Technical Report Series
38. Merriam-Webster (1987) Webster's Ninth New Collegiate Dictionary. Springfield, Massachusetts, USA
39. Hornby AS (1980) Oxford Advanced Learner's Dictionary of Current English. 3rd edn 1974. Oxford University Press
40. Commission of the European Communities (2002) European commission. Regulation (EC) No 178/2002 of the European parliament and of the council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European food safety authority and laying down procedures in matters of food safety. Official Journal of the European Commission. L31, 1–24. Commission of the European Communities
41. FAO/WHO (1998) Carbohydrates in human nutrition. FAO Food and Nutrition paper 66.http://www.fao.org/docrep/w8079e/w8079e00. htm#Contents
42. European Agency for the Evaluation of Medicinal Products (2005) Guideline for good clinical practice. http://www.emea.eu.int/pdfs/human/ich/013595en.pdf
43. The United Kingdom Good Laboratory Practice Monitoring Authority (2005) Definition of Good Laboratory Practice. http://medicines.mhra.gov.uk/ourwork/ensurequalmed/inspectlabs.htm
44. World Health Organization (2005) Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19–22 June 1946; signed on 22 July 1946 by the representatives of 61 States and entered into force on 7 April 1948. http://www.who.int/about/definition/en/.
45. Huppert FA, Baylis N (2004) Well-being: towards an integration of psychology, neurobiology and social science. Philos Trans R Soc Lond B Biol Sci 359: 1447–1451

July 10, 2003

Guidance for Industry and FDA

Interim Evidence-based Ranking System for Scientific Data

GUIDANCE

Contains Nonbinding Recommendations

Comments and suggestions regarding this document may be submitted at any time. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact the Center for Food Safety and Applied Nutrition (CFSAN) at 301-436-1450.

Additional copies are available from:
Office of Nutritional Products, Labeling, and Dietary Supplements
Division of Nutrition Programs and Labeling HFS-800
Center for Food Safety and Applied Nutrition
Food and Drug Administration
5100 Paint Branch Parkway
College Park, MD 20740
(Tel) 301-436-1450
http://www.cfsan.fda.gov/guidance.html

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Food Safety and Applied Nutrition (CFSAN)
July 2003

Guidance for Industry and FDA[1]

Interim Evidence-based Ranking System for Scientific Data

This guidance represents the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

TABLE OF CONTENTS

I. INTRODUCTION

This guidance is intended to notify the public of the Food and Drug Administration's (FDA) interim evidence-based ranking system that is a process designed to lay a foundation for a more detailed system to be used permanently. This guidance describes a process that FDA intends to use, on an interim basis, to evaluate and rank the scientific evidence in support of a substance/disease relationship that is the subject of a qualified health claim until the agency can promulgate regulations under notice-and-comment rulemaking. Based on this process, the agency will categorize the qualified health claim into one of three levels (i.e., a "B", "C", or "D" level). This guidance does not apply to unqualified health claims, which must meet the "Significant Scientific Agreement" (SSA) standard.(2)

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. BACKGROUND

This interim ranking system provides criteria to rank scientific evidence relevant to substance/disease relationships that are the subject of qualified health claims. It outlines the major concepts the agency intends to consider in guiding the scientific evaluation.

The primary purpose of this guidance is to provide petitioners with a description of the major points the agency intends to consider in evaluating supporting scientific data.

III. DISCUSSION

A. What is an Evidence-based Rating System?

An evidence-based rating system is a science-based systematic evaluation of the strength of the evidence behind a statement. In the case of health claims, it would rate the strength of the evidence behind a proposed substance/disease relationship. A large number of evidence-based rating systems are currently in use today by physicians, dietitians and other health professionals.(3) FDA has tentatively chosen to model its evidence-based rating system on that of the Institute for Clinical Systems Improvement (ICSI)(4) as adapted by the American Dietetic Association(5) with modifications specific to FDA. In making this tentative decision, FDA relied on criteria for evaluating evidence-based rating systems as reviewed and critiqued by the Agency for Healthcare Research and Quality (AHRQ).(3) FDA also found the modifications from the American Dietetic Association to be particularly useful as they considered diet and health relationships, whereas other groups focused on drug and treatment applications.

B. How are "Rate" and "Rank" Used in this System?

The terms "rate" and "rank" are not used interchangeably to describe this system. The evaluation process involves three separate **rating** systems: (1) a rating for study design; (2) a rating for study quality; and (3) a rating for the strength of the entire body of evidence. Considering all classifications from the three rating systems, a final **rank** of the scientific evidence in support of a health claim would be assigned.

C. **What are the Parts of an Evidence-based Rating System?**

In order to evaluate the level of scientific support for a proposed substance/disease relationship, the agency intends to follow a six-part procedure.

Each part of the evidence-based rating system is described below:

1. *Define the substance*[6] */disease relationship*

   A proposed relationship between a substance and a disease or health-related condition is identified. If relevant, the subgroups within the general population, for which the relationship is targeted are identified. The relationship forms the basis for selecting relevant studies and for evaluating the quality of the selected studies.

2. *Collect and submit all relevant studies*

   All relevant studies (both favorable and unfavorable) to the relationship to be tested (as defined above in C.1.) are collected and submitted. The evaluation of the proposed relationship relies primarily on human studies.

3. *Classify, and therefore rate, each study as to type of study*

   Each study would be characterized as a study design type.[7] By categorizing the study, it automatically receives an initial study "rating" based on the type of experimental design, which is independent of the quality of the study. The rating of study design is based on the principle of minimizing bias.[8] Only primary reports of data collection are rated. Reports that synthesize or reflect collections of primary reports are not considered part of the rating system although they may provide useful background information.

   a. Study Design Type One
      - Randomized, controlled intervention trials

   b. Study Design Type Two
      - Prospective observational cohort studies

   c. Study Design Type Three
      - Nonrandomized intervention trials with concurrent or historical controls
      - Case-control studies

   d. Study Design Type Four
      - Cross-sectional studies
      - Analyses of secondary disease endpoints in intervention trials
      - Case series

4. *Rate each study for quality*

   Each study would be reviewed independently and assigned a quality factor of +, Ø, - or N/A. The basis for the assignment of the quality factor is discussed below. [9]

   a. (+) means the report has adequately addressed issues of scientific quality such as inclusion/exclusion, bias, generalizability, and data collection and analysis.

   b. (Ø) means some uncertainties exist as to whether the report has adequately addressed issues of scientific quality such as inclusion/exclusion, bias, generalizability, and data collection and analysis.

   c. (-) means the report has not adequately addressed issues of scientific quality such as inclusion/exclusion, bias, generalizability, and data collection and analysis.

   d. N/A means the report is not a primary reference, therefore the quality has not been assessed, and such a reference is not considered as part of the body of evidence on which the final ranking is based. Examples of non-primary references are review articles and meta analyses.

5. *Rate the strength of the total body of evidence*

The studies are considered collectively across the evidence base in order to rate the strength of the body of evidence. The rating system is based on three factors: quantity, consistency, and relevance to disease risk reduction in the general population or target subgroup. These three factors and the final "rank" for the strength of the evidence for the "relationship" are described below.

a. Rating the body of evidence for quantity, consistency, and relevance to disease risk reduction in the general population or target subgroup.

i. *Quantity.* Considers the number of studies, the total number of individuals studied and the generalizability of the findings to the target population.

- (***) means the number of studies and the number of individuals tested (from all studies of design types one and two that are of high quality (+) combined) are sufficiently large to comfortably generalize to the target population.

- (**) means there are a sufficient number of studies and individuals tested from study design type three and higher (i.e., study design types one and two) of at least moderate quality (Ø) but uncertainties remain as to generalizability to the target population.

- (*) means that the number of studies and the number of individuals tested is insufficient to generalize to the target population.

ii. *Consistency.* Considers whether studies with both similar and different designs report similar findings.

- (***) means a sufficient number of studies of design types one and two that are of high quality (+) have consistent results. Any inconsistencies should be explained satisfactorily.

- (**) means there is a moderate consistency across all study levels.

- (*) means that the results of studies are inconsistent.

iii. *Relevance to Disease Risk Reduction in the General Population or Target Subgroup.* Considers whether or not the magnitude of the risk-reduction effect in the target population is physiologically meaningful and achievable in the general US population or a subgroup of the US general population under intake and use conditions that are appropriate for such conventional human food and human dietary supplements that would be the subject of the claim.

- (***) means that the magnitude of the effect observed in studies of design types one and two that are of high quality (+) is physiologically meaningful and achievable under intake and use conditions that are appropriate for such conventional human food and human dietary supplements that would be the subject of the claim.

- (**) means there is some suggestion from studies of design type three and higher (i.e., study design types one and two) and of moderate (Ø) to high (+) quality that the effect will be physiologically meaningful, and achievable under intake and use conditions that are appropriate for such conventional human food and human dietary supplements that would be the subject of the claim but uncertainties remain.

- (*) means that the magnitude of the effect in the studies is not likely to be physiologically meaningful or achievable under intake and use conditions that are appropriate for such conventional human food and human dietary supplements that would be the subject of the claim.

b. Ranking the Strength of the Evidence for a Health Claim

i. The first level, or highest rank of scientific evidence to support the substance/disease relationship meets the "Significant Scientific Agreement among qualified experts" standard. (For the purpose of this guidance, the first level rank is only used as a reference point. In all other respects it is outside the scope of this guidance.)

This level reflects a *high level of comfort*(10) among qualified scientists that the claimed substance/disease relationship is scientifically valid. In general, the first level ranked relationship would be considered to have a very low probability of significant new data overturning the conclusion that the relationship is valid or significantly changing the nature of the relationship. It would have high consistency with conclusions of authoritative bodies. The relationship would be based on relevant, high quality studies of mostly study design types one and two, and sufficient numbers of individuals would be tested to result in a high degree of confidence that results are relevant to the target population. Studies of different design would almost always result in similar findings, and the benefit would be physiologically meaningful and achievable under intake and use conditions that are appropriate for such conventional human food and human dietary supplements that would be the subject of the claim.

ii. The second level rank of scientific evidence to support the substance/disease relationship is the highest level for a qualified health claim, and represents a *moderate/good level of comfort* among qualified scientists that the claimed relationship is scientifically valid. Qualified experts would rank the relationship as "promising," but not definitive. The claim would be based on relevant, high to moderate quality studies of study design type three and higher (i.e., design types one and two) and sufficient numbers of individuals would be tested to result in a moderate degree of confidence that results could be extrapolated to the target population. Studies of similar or different design would generally result in similar findings and the benefit would reasonably be considered to be physiologically meaningful and achievable under intake and use conditions that are appropriate for such conventional human food and dietary supplements that would be the subject of the claim. (Note: The term "moderate/good" for the second level rank may seem ungenerous. This terminology derives from historical data evaluated by the National Academy of Sciences(11) that indicated that over time many diet/disease relationships that met this level of evidence were not necessarily sustained.)

iii. The third level rank of scientific evidence to support the substance/disease relationship is the middle level for a qualified health claim and represents a *low level of comfort* among qualified scientists that the claimed relationship is scientifically valid. It would have low consistency with statements from authoritative bodies or be ranked as "low" in terms of scientific support by qualified scientists. The relationship would be based mostly on moderate to low quality studies of study design type three, and insufficient numbers of individuals would be tested, resulting in a low degree of confidence that results could be extrapolated to the target population. Studies of different design would generally result in similar findings but uncertainties would exist. Uncertainties would also exist as to whether the benefit would be considered physiologically meaningful and achievable under intake and use conditions that are appropriate for such conventional human food and human dietary supplements that would be the subject of the claim.

iv. The fourth level, or the lowest rank of scientific evidence to support the claimed substance/disease relationship, is the lowest level for a qualified health claim and represents an *extremely low level of comfort* among qualified scientists that the claimed relationship is scientifically valid. It would have very low consistency with conclusions of authoritative bodies or be ranked very low by qualified scientists. The relationship would be based mostly on moderate to low quality studies of study design type three and insufficient numbers of individuals would be tested, resulting in a very low degree of confidence that results could be extrapolated to the target population. Studies of different design would generally result in similar findings but uncertainties would exist. There could be considerable uncertainty as to whether or not the benefit would be considered physiologically meaningful or achievable under intake and use conditions that are appropriate for such conventional human food and human dietary supplements that would be the subject of the claim. This level requires at least some credible evidence to support the relationship. There cannot be a strong body of evidence against the claim (e.g., a study or studies of high persuasiveness, quality and relevance that do not detect an effect). If that is the case, such evidence provides a

sound basis for concluding that the claim is not valid.

v. If the scientific evidence to support the substance/disease relationship is below that described as the fourth level (see above) *no claim will be appropriate.*

6. *Report the "rank"*

The result of the evidence-based rating system will be a statement describing the nature of the evidence and the rationale for linking a substance to a disease/health-related condition with a ranking as to the strength of the scientific evidence in support of that relationship. The process for arriving at the rank of the evidence to support the substance/disease relationship is illustrated in **Table 1**. The rank will be supported by:

a. A clear and transparent demonstration of which research studies were evaluated to provide the rank.

b. Evidence tables showing the rigor of the evaluation.

Table 1. Overview of the evidence-based rating system for evaluating the substance/disease relationship that is the subject of a qualified health claim.
There are six steps to evaluating the strength of the scientific evidence in support of a qualified health claim.

Step One. A proposed relationship between a substance and a disease or health-related condition is identified.
Step Two. Individual studies are identified that are pertinent to the substance/disease relationship.
Step Three. Individual studies are classified according to study design type. Different design types are graded higher than others, based on their ability to minimize bias. Thus assignment of a study design automatically provides a rating.
Step Four. Individual studies are assigned a designator of +, Ø, -, or N/A to reflect the study quality. (The general criteria for quality determination are described in this guidance).
Step Five. The strength of the scientific evidence in support of the substance/disease relationship is given a rank. This rank is determined taking into account the quantity, consistency, and relevance to disease risk reduction of the *aggregate* of the studies.
Step Six. The rank is reported.

D. What Resource Materials are Available?

1. *1. Internet-based Resource Materials*

- Agency for Healthcare Research and Quality (at http://www.ahrq.gov)
- American Dietetic Association (at http://www.eatright.org/)
- Canadian Task Force on Preventive Health Care (at http://www.ctfphc.org/)
- Center for Evidence Based Medicine (at http://www.cebm.utoronto.ca)
- Cochrane Collaboration/Cochrane Reviews (at http://www.cochrane.org)
- Evidence-based Practice Internet Resources (at http://www-hsl.mcmaster.ca/ebm/)
- Federal Judicial Center (at http://www.fjc.gov)

- Federal Trade Commission (at http://www.ftc.gov)
- FDA Food Advisory Committee. See Report of the FDA Food Advisory Committee Emerging Science Working Group at http://www.cfsan.fda.gov/~dms/faclaims.html
- FDA Guidance for Industry: Significant Scientific Agreement in the Review of Health Claims for Conventional Foods and Dietary Supplements; Availability (64 FR 71794; December 22, 1999) (see http://www.cfsan.fda.gov/guidance.html)
- Health Canada. Since their June 2000 publication of the proposed standards for health claims, proposals on two approaches to regulating health claims on foods have been published. The two approaches are: generic authorization and product-specific authorization (see http://canada.ca).
- National Coordination Centre for Health Technology Assessment (at http://www.ncchta.org/main.htm)
- National Guideline Clearinghouse (at http://www.guideline.gov)
- National Health and Medical Research Council (at http://www.health.gov.au/nhmrc/)
- National Health Service Centre for Reviews and Dissemination (http://www.york.ac.uk/inst/crd/).
- National Heart, Blood, and Lung Institute (specific information available at http://www.nhlbi.nih.gov/health/public/lung/)
- New Zealand Guidelines Group (at http://www.nzgg.org.nz/)
- Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence-based medicine: what it is and what it isn't (see http://www.cebm.net/ebm_is_isnt.asp)
- Scottish Intercollegiate Guidelines Network (at http://www.sign.ac.uk/)

2. Other Resource Materials

- Ahrens, E.H., Jr. Symposium. The evidence relating six dietary factors to the nation's health: consensus statement. Introduction. *Am. J. Clin. Nutr.* 32:2627-2631, 1979.
- Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA* 1996;276:637-39.
- Briss PA, Zaza S, Pappaioanou M, Fielding J, Wright-De Aguero L, Truman B, Hopkins D, Mullen PD, Thompson RS, Woolf SH, Carande-Kuis VG, Anderson A, Hinman AR, McQueen DV, Teutsch SM, Harris JR. Developing an evidence-based Guide to Community Preventive Services - methods. The Task Force on Community Preventive Services. *Am J Prev Med* 2000;18:35-43.
- Chalmers TC, Smith H Jr, Blackburn B et al. A method for assessing the quality of a randomized control trial. *Control Clin Trials.* 1981;2:31-49.
- Clarke M., Oxman AD. Cochrane Reviewer's Handbook 4.0. The Cochrane Collaboration; 1999.
- Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam Consultation on Meta-Analysis. *J Clin Epidemiol.* 1995;48:167-171.
- Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomized studies of health care interventions. *J Epidemiol Community Health.* 1998;52:377-384.

- Fahey T, Hyde C, Milne R, Thorogood M. The type and quality of randomized controlled trials (RCTs) published in UK public health journals. *J Public Health Med.* 1995;17:469-474.
- Falk, M. Model for a third-party review of the evidence substantiating food and dietary supplement claims. *J Nutr* 131:2219-2223, 2001.
- Goodman SN, Berlin J, Fletcher SW, Fletcher RH. Manuscript quality before and after peer review and editing at Annals of Internal Medicine. *Ann Intern Med.* 1994;121:11-21.
- Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet.* 2000;355:103-106.
- Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-based medicine working group. *JAMA.* 1993; 270:2598-2601.
- Guyatt GH, Haynes RB, Jaeschke RZ, Cood DJ, Green L, Naylor CD, Wilson MC, Richardson WS. Users' Guides to the Medical Literature: XXV. Evidence-based medicine: principles for applying the Users' Guides to patient care. Evidence-Based Medicine Working Group. *JAMA.* 2000;284:1290-1296.
- Harbour R, Miller J. A new system [Scottish Intercollegiate Guidelines Network (SIGN)] for grading recommendations in evidence based guidelines. *BMJ.* 2001;323:334-336.
- Harris RP, Helfand M, Woolf SH et al. Current methods of the U.S. Preventive Services Task Force: A review of the process. *Am J Prev Med.* 2001;20:21-35.
- Hibble, A, Kanka, D, Pencheon, D, and Pooles, F. Guidelines in general practice: the new Tower of Babel? *British Medical Journal* 317:862-863, 1998.
- Institute of Medicine. Guidelines for clinical practice: from development to use. Washington DC: National Academy Press, 1992.
- Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17:1-12.
- Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA.* 1999;282:1054-1060.
- Kaye DH, Freedman DA. Reference Guide on Statistics. In: Reference Manual on Scientific Evidence, Federal Judicial Center, 2000.
- Liberati A, Himel HN, Chalmers TC. A quality assessment of randomized control trials of primary treatment of breast cancer. *J Clin Oncol.* 1986;4:942-951.
- Lohr KN, Carey TS. Assessing 'best evidence': issues in grading the quality of studies for systematic reviews. *Joint Commission J Qual Improvement.* 1999;25:470-479.
- National Health and Medical Research Council (NHMRC). How to use the evidence: assessment and application of scientific evidence. Canberra, Australia: NHMRC,2000.
- Nowak R. Problems in clinical trials go far beyond misconduct, 264 *Science* 1538, 1994.
- Porter C, Matel JL. Are we making decisions based on evidence? *JADA.* 1998;98:404-407.
- Reisch JS, Tyson JE, Mize SG. Aid to the evaluation of therapeutic studies. *Pediatrics.* 1989;84:815-827.
- Schulz KF. Subverting randomization in controlled trials, 274 *JAMA* 1456, 1995.
- Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of

methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995;273:408-412.

- Spilker, B. Guide to Clinical Trials. Raven Press, NY, 1991. Chapter 103, Systems to Evaluate Published Data.
- Splett P. Developing and validating evidence-based guides for practice: a tool kit for dietetics professionals. Chicago: The American Dietetic Association, 2000.
- Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Systematic reviews of trials and other studies. *Health Technol Assess.* 1998;2:1-276.
- Truman BI, Smith-Akin CK, Hinman AR, Gebbie KM, Brownson R, Novic LF, Lawrence RD, Pappaioanou M, Fielding J, Evans CA, Guerra FA, Vogel-Taylor M, Mahan CS, Fullilove M, Zaza S. Developing the Guide to Community Preventive Services - overview and rationale. The Task Force on Community Preventive Services. *Am J Prev Med* 2000;18:18-26.
- West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, Lux L. Systems to Rate the Strength of Scientific Evidence. AHRQ Publication No. 02-E016, 2002.
- Zaza S, Wright-De Aguero LK, Briss PA, Truman BL, Hopkins DP, Hennessy MH, Sosin DM, Anderson L, Carande-Kulis VG, Teutsch SM, Pappaioanou M. Data collection instrument and procedure for systematic reviews in the Guide to Community Preventive Services. Task Force on Community Preventive Services. *Am J Prev Med* 2000;18:44-74.

---

(1)This guidance has been prepared by the Center for Food Safety and Applied Nutrition (CFSAN) at the U.S. Food and Drug Administration.

(2)FDA uses the term, "unqualified health claim," to refer to health claims that are or could be authorized under the Nutritional Labeling and Education Act of 1990 (NLEA) and regulations promulgated under that act, including 21 CFR 101.70.

(3)Examples of evidence-based rating systems are described and evaluated in: Agency for Healthcare Research and Quality. Systems to rate the strength of scientific evidence. Evidence Report/Technology Assessment. Number 47, 2002. The Healthcare Research And Quality Act of 1999, Part B, Title IX, Section 911(a) mandated that the Agency for Healthcare Research and Quality (AHRQ), in collaboration with experts from the public and private sectors, identify methods or systems to assess health care research results, particularly "methods or systems to rate the strength of the scientific evidence underlying health care practice, recommendations in the research literature, and technology assessments."

(4)Greer N, Mosser G, Logan G, Wagstrom Halaas G. A practical approach to evidence grading. Jt Comm. J Qual Improv. 2000; 26:700-712.

(5) The ICSI system has been adapted by the American Dietetic Association (ADA) for their evidence-based dietetics practice and, thus, the ADA modifications have addressed many of the diet/disease relationships that are also of interest to FDA. See: Myers EF, Pritchett E, Johnson EQ. Evidence-based practice guides vs. protocols: what's the difference? JADA. 2001;101:1085-1090.

(6)As defined in 21 CFR 101.14 (a)(2), the term "substance" means a specific food or component of food, regardless of whether the food is in conventional food form or a dietary suppleme nt that includes vitamins, minerals, herbs, or other similar nutritional substances.

(7)This rating system for type of study design is based on that described in Greer et al., 2000, with modifications.

(8)For example, randomization minimizes bias in that the groups are likely to be comparable except for the treatment. That is why inferences based on randomized experiments are considered more secure than inferences based on observational studies (from Kaye DH and Freedman DA. Reference Guide on Statistics. In: Reference Manual on Scientific Evidence, Federal Judicial Center, 2000.).