Improving selection of markers in nutrition research: evaluation of the criteria proposed by the ILSI Europe Marker Validation Initiative

Philip C. Calder[1], Alan Boobis[2], Deborah Braun[3], Claire L. Champ[4], Louise Dye[4], Suzanne Einöther[5], Arno Greyling[5], Christophe Matthys[6], Peter Putz[7], Suzan Wopereis[8], Jayne V. Woodside[9] and Jean-Michel Antoine[10]

[1]Faculty of Medicine, University of Southampton, Southampton SO16 6YD, UK

[2]Centre for Pharmacology & Therapeutics, Imperial College London, London W12 0NN, UK

[3]Institut Mérieux, 69002 Lyon, France

[4]School of Psychology, University of Leeds, Leeds LS2 9JT, UK

[5]Unilever R&D, 3133 AT, Vlaardingen, The Netherlands

[6]Clinical Nutrition Unit, University Hospitals Leuven & Clinical and Experimental Endocrinology, KU Leuven, 3000 Leuven, Belgium

[7]ILSI Europe a.i.s.b.l., 1200 Brussels, Belgium

[8]TNO, 3700 AJ, Zeist, The Netherlands

[9]Centre for Public Health, Queen's University Belfast, Belfast BT12 6BJ, UK

[10]Danone Research, 91767 Palaiseau Cedex, France

Corresponding author: ILSI Europe a.i.s.b.l., Avenue E. Mounier 83, Box 6, 1200 Brussels, Belgium, e-mail: publications@ilsieurope.be, tel +32 2 771 0014, fax +32 2 762 00 44

Running title: ILSI Europe Marker Initiative Follow-Up

Key words: Marker; Biomarker; Nutrition; Validation criteria; Scoring

Abstract

The conduct of high quality nutrition research requires the selection of appropriate markers as outcomes, for example as indicators of food or nutrient intake, nutritional status, health status or disease risk. Such selection requires detailed knowledge of the markers, and consideration of the factors that may influence their measurement, other than the effects of nutritional change. A framework to guide selection of markers within nutrition research studies would be a valuable tool for researchers. A multidisciplinary Expert Group set out to test criteria designed to aid the evaluation of candidate markers for their usefulness in nutrition research and subsequently to develop a scoring system for markers. The proposed criteria were tested using thirteen markers selected from a broad range of nutrition research fields. The result of this testing was a modified list of criteria and a template for evaluating a potential marker against the criteria. Subsequently, a semi-quantitative system for scoring a marker and an associated template were developed. This system will enable the evaluation and comparison of different candidate markers within the same field of nutrition research in order to identify their relative usefulness. The ranking criteria of proven, strong, medium, or low are likely to vary according to research setting, research field and the type of tool used to assess the marker and therefore the considerations for scoring need to be determined in a setting-, field- and tool-specific manner. A database of such markers, their interpretation and range of possible values would be valuable to nutrition researchers.

Introduction

A biomarker has been defined as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to an intervention"[1]. Thus, biomarkers are measurements that reflect biological processes and they can be various sorts of data, such as physiological measurements; analyses of tissues, blood or other body fluids; metabolic data; genetic data; or measurements from bio-images. In recent years new technologies have enabled the simultaneous measurement of genetic sequences, messenger RNAs, peptides, proteins, or metabolites resulting in patterns (or "signatures") as biomarkers. Biomarkers have relevance to medical practitioners and other healthcare professionals, researchers, the general public, patient sub-groups, industry, healthcare funders, regulators, and policy makers. It is important to distinguish between biomarkers, risk factors, and endpoints. Biomarkers are biological characteristics that are measured and evaluated. As a consequence, they are subject to measurement quality issues such as accuracy, precision, reliability, reproducibility, and the need for standards and quality control. Risk factors are variables that are related to an increased probability of developing a disease or injury; they may include biomarkers but also social and environmental factors. Endpoints are clinical outcomes or events. Surrogate biomarkers are substitutes for clinically meaningful endpoints and are expected to predict the effect of a therapy[2], but not all biomarkers predict risk or function as endpoints.

Biomarkers, risk factors and endpoints are all very relevant to nutrition research and are widely used. However, nutrition researchers are often interested in a broader range of exposures and outcomes. These may include food, nutrient and non-nutrient intake from the diet; behaviour in the context of food or nutrient exposure; psychological as well as physiological outcomes; and well-being. Fig. 1 depicts the relationship between dietary exposure, nutrient status, and the impact of nutrition on growth, development, behaviour, and

psychological and physiological function, which in turn influence health, wellbeing and disease risk. A concept that defines biomarkers, bio-indicators and public health indicators as types of measures in nutritional assessment has recently been introduced[3]. Hence, the term "markers" is used herein to distinguish this broader range of nutritional interests from the narrower focus upon physiological "biomarkers" (Fig. 1); note that "markers" will include "biomarkers".

The conduct of high quality nutrition research requires the selection of appropriate markers as outcomes, for example as indicators of food or nutrient intake, nutritional status, health status or disease risk. The selection of suitable markers will allow a research question to be robustly addressed, but such selection requires detailed knowledge of the markers, and consideration of the factors that may influence the measurement of these markers, other than the effects of nutritional change. A framework to guide selection of markers within nutrition research studies would be a valuable tool for researchers in the field. In this context, a key conclusion of the European Commission-funded project PASSCLAIM, coordinated by the European Branch of the International Life Sciences Institute (ILSI Europe), was that there is a lack of adequate markers in nutrition sciences and that there is a high need for such markers[4]. ILSI Europe therefore launched an activity, "Marker Initiative on Nutrition Research", with the aim of identifying and reviewing criteria for validation of markers. It was envisaged that this would be a multi-step process, as illustrated in Fig. 2. The first step was the identification of those criteria that have been used to assess a broad range of markers in nutrition research. This was followed by a Workshop, "Obtaining consensus on the criteria for evaluating markers in nutrition research", held in June 2012 in Lisbon, Portugal, comprising Step 2 of the process (see Fig. 2). During the Workshop, participants established a preliminary list of consensus criteria for marker evaluation for nutrition research[5]:

1) the marker should be validated according to recognised methods;

2) the marker should reflect an endpoint (there should be a significant association between the marker and an endpoint in a target population and the marker should change consistently with a change in the endpoint);

3) the marker must respond to a dietary intervention.

The next step in the process (Step 3), the current activity, was to assess the use of these criteria, using a range of different possible markers reflecting the breadth of nutrition research possibilities in order to test whether the criteria were fit-for-purpose, and, if not, to propose alternatives. The current activity was performed by a multi-disciplinary Expert Group, members of which discussed all aspects under consideration until consensus was reached. This article conveys the result of those discussions. One outcome of the current activity is a revised list of criteria, incorporated into a template. A second role of the current activity was to consider the development of methods for scoring markers against the pre-specified criteria, and to develop a template for this purpose. Review of "nutritional (bio)markers" themselves was outside the remit of the expert group, the overall aim of which was to (re)consider the process by which such markers can be evaluated.

Testing the proposed criteria

It was considered that the best way to test the criteria was to complete a template based on the criteria established by de Vries et al.[5] (see Table 1) using examples of markers that reflect:

1) a broad range of interests in nutrition research (see Fig. 1);
2) the use of different tools, including both questionnaires and laboratory tests;
3) both long established and newer markers and tools;
4) commonly used and not commonly used markers.

Hence, the markers selected are not necessarily well validated or widely accepted. Markers covering the fields of nutrient intake, nutrient status, physiological function, metabolism, cognitive function, and disease risk were all evaluated (Table 2).

Individual members of the multidisciplinary Expert Group completed the draft template (Table 1) using the criteria proposed in Step 2 (de Vries et al.[5]) (henceforth "proposed criteria", see Fig. 2) for each of the markers selected and then the completed template was discussed amongst all members and modifications made until consensus on the utility of the criteria for each marker was reached. One completed template is included (Table 3), while the completed template for each of the 13 markers is included as Supplementary Material.

Refining the criteria and developing a new evaluation template

As experts evaluated the proposed criteria and during the subsequent discussions, a number of pertinent points emerged regarding the ease of use, utility and relevance of the criteria and also the exact wording used to describe/define criteria. Although several of the core components that form part of the criteria are clearly defined (e.g. sensitivity, specificity), others are not (e.g. robustness) making it difficult to address these less well defined criteria. Furthermore, in the absence of clear definitions, different individuals interpret the meaning of these terms or criteria differently, resulting in a less robust and less reproducible (from individual to individual) evaluation. It was recognised that, despite that fact that some markers are widely used, they fail to meet some of the criteria; for example in some cases assays may be poorly standardised. Furthermore, some measures are often used as endpoints themselves rather than as markers of other endpoints (outcomes); for example verbal memory, a marker of cognitive function, is often reported as an endpoint in its own right. Also, some markers have remained in use over a very long period of time, perhaps because they become validated or well accepted or, in some cases, because they are easy to use. On

149 the other hand, some markers cease to be used after a period of time while new markers can
150 emerge. Thus, there is a certain level of "turnover" of markers. This has been hastened by the
151 emergence of new technologies, typically "omics"-based, that have enabled the simultaneous
152 measurement of clusters or patterns of markers. Some of these seem likely to replace existing
153 single-measurement markers, although validation of the patterns, access to the technology
154 and cost remain barriers. An example of how the marker field is developing is the proposal of
155 a composite biomarker called the "vascular health index" based on the integration of a
156 meticulously selected cluster of biomarkers all related to vascular health measured with
157 different types of analytical techniques[6].
158
159 Many markers are evaluated in a static setting, for example in fasting blood samples. This
160 separates the sample and its component markers from the reality of human physiology, which
161 is the need to respond appropriately to "daily stressors" which may be metabolic, immune,
162 physical (e.g. exercise, temperature), or psychological. Thus, it may be desirable to include a
163 challenge test in a study protocol and to evaluate the dynamic change in the marker in
164 response to the challenge. Some of the examples of markers used to test the criteria do
165 include a challenge test (e.g. response to vaccination); again it seems that in the future more
166 studies will incorporate challenge models to simulate events which could occur in the natural
167 environment[7,8]. One area where challenge models have become widely used in recent years
168 in order to study the dynamic change in a marker is the evaluation of inflammation. For
169 example, both high fat and high carbohydrate meals induce an acute elevation in a number of
170 biomarkers of inflammation[9] and such challenges have been used to assess the effect of
171 including fibre[9] or vitamin C[10] in the meal on the acute inflammatory response that is
172 elicited. Exposure of the skin to ultraviolet irradiation induces inflammation and controlled
173 exposures have been used to assess the effects of including omega-3 fatty acids[11] or green
174 tea catechins[12] in the diet on a range of biomarkers of inflammation. Intramuscular injection

of bacterial endotoxin has been used to assess the effect of dietary omega-3 fatty acids on inflammation[13].

It also emerged that a marker may not be equally useful across different applications. For example, a marker that is informative in the controlled setting of a small intervention study may be much less informative, or even unfeasible, in the setting of a large observational study.

The proposed criteria included "Must respond to a dietary intervention". However, upon further discussion it was considered that whether a measure (i.e. a marker) is sensitive to nutrition does not make it a better or worse marker. This will depend on what the marker is designed to measure. Further, when considering whether a marker is influenced by a dietary or other intervention, then the extent of the effect seen (or foreseen) needs to be taken into account. This poses a challenge because studies are typically powered to show a statistically significant change in the marker being used. Even if that marker has an association with an endpoint, a statistically significant change in the marker may not be of clinical significance or even biological significance. Conversely a change that is clinically significant may not be statistically significant in a study setting. Thus, it would seem prudent when planning a nutrition study to consider both clinical and statistical significance of the change being sought. Such considerations of study design, including effect size, are discussed elsewhere[14]. It was concluded that the two criteria listed in the section "Reflect/mark an endpoint" were essentially addressing the same point, that a relationship exists between the marker and an endpoint of interest. It was difficult for experts to adequately complete the section on "Analytical aspects" because the different criteria asked about in this section were not well separated. Thus, by completing the draft template, based upon the proposed criteria, a number of the components of the proposed criteria were identified for change or improvement. It was

201 also identified that providing information on normal values or ranges in different population
202 sub-groups and thresholds used to make different conclusions would be very valuable and
203 was not explicitly requested in the draft template. It was also felt that sections in a new
204 template to add other relevant information, for example to record inconsistencies in the
205 literature, to make a clear conclusion about the usefulness of the marker under consideration,
206 including any important limitations, and to record references used would all be valuable.
207
208 The above considerations led the Expert Group to conclude that the proposed criteria could
209 be improved upon and therefore the draft template (Table 1) was revised to produce a new
210 template reflecting these improvements (Table 4). This retains the general features of the
211 proposed criteria (as described in the draft template), but the template is formatted in a way
212 that is easy for the end-user to complete with a clearer indication of the nature of the
213 information that is required for each section. The section "Methodological aspects"
214 (previously termed "Analytical aspects") explicitly separates the most important components
215 (validation; sensitivity; specificity; technical aspects other than sensitivity and specificity;
216 biological variation), providing an opportunity to consider these individually. The section
217 "Reflects the biological purpose of the marker" (previously "Reflects/marks an endpoint")
218 combines the two previous criteria (Significant association between marker and endpoint in a
219 target population and Marker changes consistently with a change in the endpoint) into a
220 single reworded criterion (A change in the marker is linked with a change in the endpoint in
221 one or more target population(s)). This is because the two previous criteria address the same
222 point, both stating that a relationship exists between the marker and an endpoint of interest.
223 The section "Relevance to nutrition research" (previously termed "Must respond to a dietary
224 intervention") expands upon and presents a change in focus from the proposed criteria. Now
225 information on the normal range of values can be entered and the requirement that the marker
226 must respond to a dietary intervention is replaced by a question seeking the evidence that

nutrition can influence the marker and, if so, the extent of the reported effect. The reason for this change in focus is that whether a measure is sensitive to nutrition or not does not make it a better or worse marker, although it may make it more or less attractive to researchers and other stakeholders. Finally, in this section a question about other factors that might affect the marker is now included. A section "Other relevant information" is included and there are cells for "Conclusion" and to record "References". It is considered that these changes will make the criteria and associated template more useful and more robust.

It is anticipated that once a particular marker has been assessed using the criteria and the associated template, the information will be used as the basis for scoring the marker in order to determine its usefulness as a research tool. Such scoring requires a suitable methodologic approach (see next section).

Towards developing a marker scoring system

Examples of scoring systems may be found in Albers et al.[15,16] where markers of immune function were evaluated by scoring them against a range of predetermined criteria; in Albers et al.[16] these related to clinical relevance, biological sensitivity, and feasibility. Table 5 is the generic marker scoring template now proposed. Researchers would score any marker under consideration according to the different criteria listed in the upper section of Table 4 as proven (+++), strong (++), medium (+), or low (0). Additionally, an arbitrary marker score would be based on subjective expert judgement on the usefulness of a marker based on carefully considered evaluation of individual criteria. This would enable researchers to evaluate and to compare different candidate markers within the same field of nutrition research in order to identify their relative usefulness. The criteria for a ranking of proven, strong, medium, or low are likely to vary according to the research setting (e.g. epidemiology, intervention, mechanistic investigation), the research field (e.g. immune

function, cognitive function, metabolism and metabolic dysfunction), and the type of tool used to assess the marker and therefore the ranking criteria need to be determined in a setting-, field- and tool-specific manner. Examples of such criteria and their use in evaluating immune markers may be found in Albers et al.[16]. It is expected that researchers would develop scoring definitions and then score and rank potential markers as part of study planning.

Summary and conclusions

An Expert Group set out to test proposed criteria (see Table 1) designed to aid the evaluation of candidate markers for their usefulness in nutrition research and subsequently to develop a scoring system for markers. The criteria were tested using a total of 13 markers selected from a breadth of fields of nutrition research (Table 3 and Supplementary Material). The result of this testing was a modified list of criteria and a template (see Table 4). It is considered that these changes will make marker assessment easier and more robust. Subsequently a system for scoring a marker and an associated template were developed (Table 5). This system would enable researchers to evaluate and to compare different candidate markers within the same field of nutrition research in order to identify their relative usefulness. The ranking criteria of proven, strong, medium, or low are likely to vary according to research setting, research field and the type of tool used to assess the marker and therefore the criteria need to be determined in a setting-, field- and tool-specific manner. Examples of such ranking criteria for immune function markers may be found in Albers et al.[16]. It is anticipated that defining the scoring system and then using this to score possible markers would be done by researchers as a part of their study planning. These activities and the development of the templates described in Tables 4 and 5 complete Step 3 of ILSI Europe's marker initiative program (see Fig. 2). The next step is to use the evaluation criteria and scoring system to evaluate markers. It is anticipated that ILSI Europe will hold an open access "library" of

completed evaluations, using the templates developed in this activity that will be available to the nutrition community for use, comment, modification, and updating. Besides applying the evaluation tool in study planning, researchers would complete templates with scenarios of marker applications to populate such a library with evaluated markers from various fields of nutrition research. It is important to note that many of the markers considered here and in the future are of interest to research communities beyond the field of nutrition and, as such, the library of ILSI Europe marker evaluations will be a valuable resource for a wide research community and for other relevant stakeholders (e.g. industry, regulators, medical practitioners).

Financial Support

This work was conducted by an expert group of the European branch of the International Life Science Institute (ILSI Europe). The expert group received funding from the ILSI Europe Functional Foods Task Force. Members of this task force are listed on the ILSI Europe website at www.ilsi.eu. The opinions expressed herein and the conclusions of this publication are those of the authors and do not necessarily represent the views of ILSI Europe, nor those of its member companies and the authors affiliations.

Conflict of Interests

Deborah Braun is an employee of Institut Mérieux, Suzanne Einöther and Arno Greyling are employees of Unilever, Jean-Michel Antoine is an employee of Danone and Peter Putz is an employee of ILSI Europe. The remaining authors declare no conflicts of interest.

Authorship

All authors contributed to discussions and had input into writing the article. P.C.C. had responsibility for producing the final version of the article.

Figure Captions

Fig. 1. Contexts of markers in nutrition research. There is a relationship between dietary exposure, nutrient status, and the impact of nutrition on growth, development, behaviour, and psychological and physiological function, which in turn influence health, wellbeing and disease risk. Nutrition research requires validated markers for each of these levels. Note that the same measure may serve as both a marker and an outcome, depending upon the context.

Fig. 2. The ILSI Europe Marker Initiative on Nutrition Research: A stepwise approach towards criteria for the evaluation of markers in different fields of nutrition research

316 **Table 1.** Template to aid the evaluation of candidate markers for their usefulness in nutrition
317 research according to previous step 2 of the Marker Initiative on Nutrition Research (de Vries
318 et al., 2013[5])

Criteria identified at the workshop	Comments
1) **Analytical aspects** i. Method should be validated according to recognised guidelines. ii. This should include for example: robustness, appropriate analytical sensitivity and specificity, reproducibility accuracy, quality assurance, standardization, traceability, stability (quality of the sample). These guidelines can be specific per marker, it apply equally well to biological markers, imaging, questionnaires, etc.	
2) **Reflect/mark an endpoint** i. Significant association between marker and endpoint in a target population ii. Marker changes consistently with a change in the endpoint	
3) **Must respond to a dietary intervention** i. Any dietary intervention should induce a meaningful change in the marker (meaningful refers to 2.ii) ii. Lifestyle changes (which may include changes in the diet) may also induce a meaningful (meaningful refers to 2.ii) change in the marker	

319
320

321 **Table 2.** Markers used to assess the proposed criteria according to their specific field of
322 application

Specific field of application	Marker
Nutritional epidemiology	Vitamin C intake determined by food frequency questionnaire
Nutritional epidemiology	Serum/plasma vitamin B12 as a status marker
Immune function	Response to vaccination
Cognitive function	Verbal memory
Cognitive function	Sustained attention
Cardiovascular diseases/Chronic inflammation	C-reactive protein concentration in plasma/serum
Cardiovascular diseases/vascular function	Flow mediated dilation
Cardiovascular diseases/oxidative stress	F2-isoprostane concentrations in urine
Cardiovascular diseases – blood pressure	Blood pressure
Metabolism and metabolic dysfunction	Branched chain amino acids and their derivatives in plasma
Metabolism and metabolic dysfunction	FADS1 genetic polymorphisms
Intestinal barrier function/Intestinal permeability	Lactulose/mannitol ratio in urine
Energetics/Obesity	Energy expenditure measured by double labelled water

323

Table 3. Example of a completed template: use of response to vaccination as a marker of immune competence. Note that references cited in this table are listed in the Supplementary Material.

Criteria to evaluate markers	Comments
1) **Analytical aspects** i. Method should be validated according to recognised guidelines. ii. This should include for example: robustness, appropriate analytical sensitivity and specificity, reproducibility, accuracy, quality assurance, standardisation, traceability, stability (quality of the sample), analytical variation, biological variation (these guidelines can be specific per marker; it applies equally well to biological markers, imaging, questionnaires, etc.)	Vaccination is a means of exposing the immune system to one or more antigens in a standardised and controlled manner. In response to vaccination, the host mounts an immune reaction that culminates in the production of anti-vaccine antibodies. The use of vaccination in an experimental setting (e.g. in a study of a nutritional intervention) involves vaccinating a participant (usually intramuscularly) with a commercial clinically used vaccine (e.g. influenza, tetanus, pneumococcus,) and obtaining blood samples at specific time points thereafter. Anti-vaccine antibodies are measured in serum or plasma prepared from the blood. The serum/plasma needs to be stored frozen (minus 80°C) until antibodies are measured. Accredited laboratories can be used to measure responses to some vaccines (e.g. WHO accredited laboratories for anti-influenza vaccine antibodies). Such laboratories use validated methodology that is recognised by WHO and other authorities. This may not be the case for all anti-vaccine antibody measurements. Where there is an accredited laboratory many of the analytical aspects are of the highest quality. For some vaccines there are definitions of seroprotection and seroconversion, although these definitions can differ between countries. Seroprotection means that the individual has a high probability of being protected and it is defined as having an antibody titre (level) above a particular threshold. For seasonal influenza vaccination, seroprotection is defined as having an antibody level of 40 haemagglutination units/ml or higher, while for diphtheria and tetanus it is defined as having an antibody level of 0.1 IU/ml or higher. Seroconversion is commonly defined as having at least a 4-fold increase in antibody levels following vaccination. Discussions of the relation between seroprotection, seroconversion and clinical protection may be found in the literature. Note that an individual can show seroconversion but still not be seroprotected and that, conversely, reaching the threshold for seroprotection may require less than a 4-fold increase in antibody levels (i.e. may not require seroconversion). Within a population there is substantial variation in the antibody response to many vaccines. The tables below show unpublished data (AL Lomax & PC Calder) from a study of 43 healthy humans aged 40 to 65 years vaccinated with the 2008/2009 seasonal influenza vaccine. The data are based upon serum antibody titres 4 weeks post-vaccination. The seasonal influenza vaccine contains subunits of three different viral strains so that three separate antibody responses are measured.

Table S1. Percentage of subjects who seroconverted and became seroprotected from a study of 43 healthy humans aged 40 to 65 years vaccinated with the 2008/2009 seasonal influenza vaccine (unpublished data)

Viral antigen type	% of subjects who seroconverted	% of subjects who became seroprotected	Antibody titre (IU/mL)		
			Lowest	Median	Highest
HAH1	79.1	72.1	5	80	15360
HAH3	79.1	81.4	5	320	20480
HAB	59.1	60.4	5	60	960

% who seroprotected				% who seroconverted			
To none of the viral antigens	To only one of the viral antigens	To any two of the viral antigens	To all three of the viral antigens	To none of the viral antigens	To only one of the viral antigens	To any two of the viral antigens	To all three of the viral antigens
4.7	18.6	34.9	41.9	2.3	20.9	34.8	41.9

These data show large variations in antibody response between individuals (i.e. within a population) and also that there is variation in response to several antigens administered together within an individual. The ability of vaccinations to initiate a robust host immune response, and so to produce clinical protection, is recognised to be poorer in the elderly (see below), the frail, the malnourished and those with certain chronic diseases.

For clinical protection against some diseases the same vaccine can be used unchanged over many years. However because of the rapid mutation rate of seasonal influenza viruses, the exact make-up of the seasonal influenza vaccine changes from year to year. The three stains that have been used in the vaccine over the years may be found on many websites (e.g. http://www.who.int/influenza/vaccines/virus/recommendations/en/). Some vaccinations (e.g. polio) give life-long protection, others (e.g. tetanus) give shorter protection.

For a primary antibody response, the subject cannot have received the same vaccine previously. Administration of a vaccine to a person who has received it already can induce a secondary antibody response, which may be different in kinetics and vigour from the primary response.

Responses to vaccination may be modified by many factors including age and health status. For example, the success of seasonal influenza vaccine is much less in people aged over 65 years than in middle aged adults. Goodwin *et al.* (2006) noted that in young healthy adults the seasonal influenza vaccine provides a protective

	clinical efficacy in 70 to 90% of cases, which is reduced to only 17 to 53% in elderly individuals. This reflects a general decline in cell-mediated immunity that occurs, to varying extents, with ageing; this is termed immunosenescence.
2) **Reflect/mark an endpoint** i. Significant association between marker and endpoint in a target population ii. Marker changes consistently with a change in the endpoint	The production of antibodies in response to vaccination represents an integrated read-out of the immune response – it will have required the activity of antigen processing and presenting cells, T cells and B cells. It is considered to be superior to any individual laboratory-based immune marker and is one of the few immune biomarkers considered to be of high value in human nutrition studies. Response to vaccination can be defined by seroprotection and seroconversion (described in the previous section) and is considered to be related to clinical outcome (i.e. protection from the infective agent), although this can be poorly defined. For seasonal influenza vaccination seroprotection and seroconversion thresholds are defined by WHO (see previous section) – in clinical practice these are often not met (e.g. in the elderly) but neither the clinician nor the patient is aware of this. Such a failure may allow susceptibility to the infectious agent, so in this sense a poor response can increase risk of poor clinical outcome (i.e. infection and its severity) while a good response can decrease risk of poor clinical outcome (i.e. infection and its severity).
3) **Must respond to a dietary intervention** i. Any dietary intervention should induce a meaningful change in the marker (meaningful refers to 2.ii) ii. Lifestyle changes (which may include changes in the diet) may also include a meaningful (meaningful refers to 2.ii) change in the marker	Response to vaccination may be used in epidemiology studies to investigate the association between the immune response and a future clinical outcome (prospective study) or between the intake of foods or nutrients and the immune response (cross-sectional study) or in intervention trials of dietary or nutrient modifications. The response to vaccination, based upon antibody titres measured at an appropriate time point after vaccination, may be represented in several different ways: antibody titre concentrations; change in antibody titre concentrations from the pre-vaccinated state; fold change in antibody titre concentrations from the pre-vaccinated state; % of individuals seroprotected; % of individuals seroconverted; % of individuals seroconverted and seroprotected (see previous section). Each of these may be a valid marker of immune response, although the clinical meaning of the outcomes may be different. For example, in a controlled trial it is possible that no subjects seroconvert or become seroprotected but that there is still a statistically significant effect of an intervention on antibody titres compared with a control group. Conversely, it is possible that all subjects in a study seroconvert and become seroprotected but that there is still a statistically significant effect of an intervention on antibody titres compared with a control group. In both of these scenarios it might be interpreted that there is an improvement in immune response, although the clinical meaning of this improvement may be different. Some studies have shown improved response to vaccination with a dietary intervention, while others have not. Studies involving prebiotics and response to vaccination were included in the review by Lomax and Calder (2009a), while studies of probiotics were reviewed by Lomax and Calder (2009b) and by Maidens *et al.* (2013). Examples of studies where improvements in response to seasonal influenza vaccination have been demonstrated include Boge *et al.* (2009), Langkamp-Henken *et al.* (2006) and Langkamp-Henken *et al.* (2004).

	Boge *et al.* (2009) reported two controlled studies of oral *Lactobacillus casei* DN-114 001 administration for 7 weeks (study 1) and 13 weeks (study 2) in elderly subjects (> 70 years of age; mean age ~ 84 years) who received the seasonal influenza vaccine after 4 weeks. Study 1 used the 2005/2006 seasonal influenza vaccine and identified 40 to 50% seroprotection and 20 to 50% seroconversion (depending upon viral strain) in the control group 3 weeks post-vaccination, in keeping with the relatively poor response to vaccination in the elderly. This study, which involved approx. 40 subjects per group (i.e. control and *L. casei*), showed non-significant trends towards higher antibody titres and increased seroprotection and seroconversion with *L. casei*. Study 2 was larger (~110 subjects in each group) and measured antibody titres 3, 6 and 9 weeks post-vaccination with the 2006/2007 vaccine. The study revealed higher antibody titres to all 3 vaccine strains at all 3 post-vaccination time points, although these were significant only for antibodies against the B strain (at all 3 time points); typically antibody titres were 60 HIU/ml in the control group and 100 in the *L. casei* group. Seroconversion to the B strain was low in the control group (~10% at 3 weeks) but was significantly increased by *L. casei* at all times points assessed (~20% at 3 weeks). Seroconversion at 5 months post-vaccination was higher to the B strain (~7% vs ~2%) and to the H3N2 strain (~20% vs ~10%) with *L. casei*. Langkamp-Henken *et al*. (2004) reported a controlled study of 26 weeks administration of an oral nutritional formula providing vitamins E and C, beta-carotene, B vitamins, Se, Zn, structured triglyceride and the prebiotic fructooligosaccharide. Subjects were older adults (> 65 years; mean age ~ 82 years) who received the 1999/2000 seasonal influenza vaccine after 2 weeks. Antibody titres to the H3N2 and B strains were not different between groups at 6 weeks or 24 weeks post-vaccination. In contrast, 6 weeks post vaccination antibody titres to the H1N1 strain were significantly higher in the formula group (~200 vs ~100). Seroprotection to each of the three viral strains was not different between groups at 6 weeks post-vaccination. Seroconversion to the H1N1 strain was higher in the formula group (87% vs 41%) but there was no difference between groups in seroconversion to the H3N2 or B strains. In the formula group 87% of subjects seroconverted and seroprotected to the H1N1 strain which was significantly higher than in the control group (35%). This study also reported cold and influenza symptoms over the study period: days with symptoms per subject were lower in the formula group (median 0 vs 3). Langkamp-Henken *et al*. (2006) reported a controlled study of 10 weeks administration of the same oral nutritional formula in older adults (> 65 years; mean age ~ 85 years) who received the 2002/2003 seasonal influenza vaccine after 4 weeks. The study identified ~45% seroconversion to the H1N1 strain and ~35% to the H3N2 strain with no effect of the nutritional formula. Seroconversion to the B strain was ~50% in the control group and 64% in the formula group (P = 0.09). Seroprotection to the H1N1 strain was higher in the formula group (44% vs 29%), and tended to be higher to the H3N2 strain (97% vs 89%) but wasn't different to the B strain (95% vs 94%). Antibody titres to any strain were not different between groups.
Conclusions	Response to vaccination, assessed as anti-vaccine antibodies in serum or plasma, can be used as a biomarker of immune function in both epidemiological and intervention studies. However, there is high variability in response

	between subjects.

327 **Table 4.** Refined template to aid the evaluation of candidate markers for their usefulness in
328 nutrition research

Criteria to evaluate markers		**Comments (white boxes to be completed)**
Specific field and related marker		
SCORING CRITERIA	**Methodological aspects (excluding study design)** Relevance of criteria can differ between different types and applications of markers	
	Method should be validated according to recognised guidelines (please cite)	
	Appropriate* sensitivity	
	Appropriate* specificity	
	Reproducibility, accuracy, standardisation, stability (quality of the sample) and technical variation	
	Biological variation	
	Reflect/mark the study objective	
	A change in the marker is linked with a change in the endpoint in one or more target population(s)	
	Method should be validated according to recognised guidelines (please cite)	
ADDITIONAL INFORMATION	**Relevance to nutrition research**	
	What is considered as a normal range for healthy people? What is a significant change (consider both biological and statistical)? (might vary for different applications e.g. epidemiological studies vs. individual level)	
	Is there evidence that nutrition influences the marker? If so, what is the size of the effect reported?	
	Which other factors also have an effect on the marker? (if any)	
	Other relevant information	
	Are there experimental data where dietary intervention has not resulted in an anticipated change?	
Conclusions		
References		

*Appropriate is used here to indicate that the required sensitivity and specificity of measurement may differ between study contexts, for example between a large epidemiological study and a much smaller randomised controlled trial.

Table 5. Generic scoring system to evaluate and compare candidate markers within the same field of nutrition research

	Methodological aspects (excluding study design)				Reflect/mark the study objective
Levels	**Appropriate sensitivity**	**Appropriate specificity**	**Reproducibility, accuracy, standardisation, stability (quality of the sample) and technical variation**	**Biological variation**	**A change in the marker is linked with a change in the endpoint in one or more target population(s)**
Proven (+++)	False negative rate (β) is well documented to be less than 1%	False positive rate (α) is well documented to be less than 1%	Marker is highly reproducible (ICC>0.9) and accurate (less than 1% deviation from "correct" value), assay is highly standardised, sample is stable or can easily be made stable	Minimal variation and relevant effects highly superior to variation: effects likely to be observed between groups of tens of people	Generally accepted marker (marker changes consistently linked with a change in the endpoint)
Strong (++)	False negative rate (β) is well documented to be less than 5%	False positive rate (α) is well documented to be less than 5%	Marker is reproducible and accurate enough to detect biologically meaningful changes, assay is standardised, sample is stable or can be made stable	High variation explainable e.g. circadian cycle, age, sex, BMI, ethnicity and genotype) and possible to correct it and relevant effects reproducibly superior to variation: effects likely to be observed between groups of fifties to hundreds of people	Described as a cause and effect relationship, but not (yet) generally accepted as marker, due to a lack of (specific) studies
Medium (+)	False negative rate (β) is well documented to be less than 10%	False positive rate (α) is well documented to be less than 10%	Marker is reproducible and accurate enough for specific applications, assay is somewhat standardised and sample needs to be extensively processed or analysed fast	High variation explainable (e.g. circadian cycle, age, sex, BMI, ethnicity and genotype) and possible to correct it and relevant effects reproducibly close to variation: effects may be observed between groups of fifties to hundreds of people	Body of evidence suggesting correlation, but cause and effect not established
Low (0)	False negative rate (β) is not well documented or shown to be at least 10%	False positive rate (α) is not well documented or shown to be at least 10%	Reproducibility and accuracy of the marker, standardisation of the assay and stability of the sample are either poor or not properly documented	High and unexplained variation in a short time span and relevant effects likely to be observed between groups of thousands of people	Plausible hypothesis, in use as an exploratory marker, but no substantial body of evidence (yet)

Reference List

1. Ball JR and Micheel CM (editors) (2010) *Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease*. Washington: National Academies Press.

2. Temple R (1999) Are surrogate markers adequate to assess cardiovascular disease drugs? *JAMA* **282**, 790-795.

3. Raiten D & Combs G (2015) Directions in nutritional assessment: Biomarkers and bio-indicators-providing clarity in the face of complexity. *Sight and Life Newsletter* **29**, 39-44.

4. Aggett PJ, Antoine JM, Asp NG*, et al.* (2005) Passclaim, consensus on criteria. *Eur J Nutr* **44**, 1-30.

5. de Vries J, Antoine JM, Burzykowski T*, et al.* (2013) Markers for nutrition studies: review of criteria for the evaluation of markers. *Eur J Nutr* **52**, 1685-1699.

6. Weseler AR & Bast A (2012) Pleiotropic-acting nutrients require integrative investigational approaches: the example of flavonoids. *J Agric Food Chem* **60**, 8941-8946.

7. van Ommen B, van der Greef J, Ordovas JM*, et al.* (2014) Phenotypic flexibility as key factor in the human nutrition and health relationship. *Genes Nutr* **9**, 1-9.

8. Stroeve JH, van Wietmarschen H, Kremer BH*, et al.* (2015) Phenotypic flexibility as a measure of health: the optimal nutritional stress response test. *Genes Nutr* **10**, 1-21.

9. Esposito K, Nappo F, Giugliano F, Di Palo C, Ciotola M, Barbieri M, *et al.* (2003) Meal modulation of circulating interleukin 18 and adiponectin concentrations in healthy subjects and in patients with type 2 diabetes mellitus. *Am J Clin Nutr* **78**, 1135-1140.

10. Peluso I, Villano DV, Roberts SA, Cesqui E, Raguzzini A, Borges G, *et al.* (2014) Consumption of mixed fruit-juice drink and vitamin C reduces postprandial stress induced by a high fat meal in healthy overweight subjects. *Curr Pharm Des* **20**, 1020-1024.

11. Pilkington SM, Massey KA, Bennett SP, Al-Aasswad NM, Roshdy K, Gibbs NK, *et al.* (2013) Randomized controlled trial of oral omega-3 PUFA in solar-simulated radiation-induced suppression of human cutaneous immune responses. *Am J Clin Nutr* **97**, 646-652.

12. Rhodes LE, Darby G, Massey KA, Clarke KA, Dew TP, Farrar MD, *et al.* (2013) Oral green tea catechin metabolites are incorporated into human skin and protect against UV radiation-induced cutaneous inflammation in association with reduced production of pro-inflammatory eicosanoid 12-hydroxyeicosatetraenoic acid. *Brit J Nutr* **110**, 891-900.

13. Michaeli B, Berger MM, Revelly JP, Tappy L, Chioléro R (2007) Effects of fish oil on the neuro-endocrine responses to an endotoxin challenge in healthy volunteers. *Clin Nutr* **26**, 70-77.

14. Welch RW, Antoine JM, Berta JL*, et al.* (2011) Guidelines for the design, conduct and reporting of human intervention studies to evaluate the health benefits of foods. *Br J Nutr* **106**, S3-S15.

15. Albers R, Antoine JM, Bourdet-Sicard R*, et al.* (2005) Markers to measure immunomodulation in human nutrition intervention studies. *Br J Nutr* **94**, 452-481.

16. Albers R, Bourdet-Sicard R, Braun D*, et al.* (2013) Monitoring immune modulation by nutrition in the general population: identifying and substantiating effects on human health. *Br J Nutr* **110**, S1-S30.