

Documentation of Functional and Clinical Effects of Infant Nutrition: Setting the Scene for COMMENT

Berthold Koletzko^a Hania Szajewska^b Margaret Ashwell^c Raanan Shamir^d Peter Aggett^e
Kurt Baerlocher^f Paul Noakes^g Christian Braegger^h Philip Calder^g Cristina Campoy Folgosoⁱ
Virginie Colomb^j Tamás Decsi^k Magnus Domellöf^l Christophe Dupont^j Mary Fewtrell^m
Johannes B. van Goudoeverⁿ Kim F. Michaelsen^o Walter Mihatsch^p Alfredo Guarino^q
Sibylle Koletzko^a Jacques Rigo^r Dominique Turck^s Jan Taminiauⁿ
on behalf of the Early Nutrition Academy and European Society for Paediatric Gastroenterology,
Hepatology and Nutrition Committee on Nutrition – Consensus Group on Outcome
Measures Made in Paediatric Enteral Nutrition Clinical Trials

^aDr. von Hauner Children's Hospital, University of Munich Medical Centre, Munich, Germany; ^bThe Medical University of Warsaw, Warsaw, Poland; ^cAshwell Associates (Europe) Ltd., Ashwell, UK; ^dSchneider Children's Medical Center of Israel, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; ^eSchool of Medicine and Health, Lancaster University, Lancaster, UK; ^fOstschweizerisches Kinderspital, St. Gallen, Switzerland; ^gSchool of Medicine, University of Southampton, Southampton, UK; ^hKinderspital, University of Zurich, Zurich, Switzerland; ⁱDepartment of Paediatrics, University of Granada, Granada, Spain; ^jHôpital Necker – Enfants Malades, Paris, France; ^kDepartment of Paediatrics, University of Pecs, Pecs, Hungary; ^lDepartment of Paediatrics, University of Umea, Umea, Sweden; ^mInstitute of Child Health, University College London, London, UK; ⁿDepartment of Paediatrics, Emma Children's Hospital – Academic Medical Centre and VU University Medical Centre Amsterdam, Amsterdam, The Netherlands; ^oInstitute of Human Nutrition, University of Copenhagen, Copenhagen, Denmark; ^pEvangelisches Diakoniekrankenhaus, Schwäbisch Hall, Germany; ^qDepartment of Paediatrics, University of Naples, Naples, Italy; ^rDepartment of Paediatrics, University of Liege, Liege, Belgium; ^sJeanne de Flandre Children's Hospital, University of Lille, Lille, France

Key Words

Consensus Group on Outcome Measures Made in Paediatric Enteral Nutrition Clinical Trials · European Society for Paediatric Gastroenterology, Hepatology and Nutrition · Early Nutrition Academy · Functional effects · Nutritional interventions · Infants · Young children · Clinical trials · Markers · Outcome measures · Allergy · Growth · Diarrhoea · Core data

Abstract

The Early Nutrition Academy and the Child Health Foundation, in collaboration with the Committee on Nutrition, European Society for Paediatric Gastroenterology, Hepatology and Nutrition, held a workshop in March 2011 to explore guidance on acquiring evidence on the effects of nutritional interventions in infants and young children. The four objectives were to (1) provide guidance on the quality and quantity of evidence needed to justify conclusions on functional and clinical effects of nutrition in infants and young children aged <3 years; (2) agree on a range of outcome measures

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2012 S. Karger AG, Basel
0250–6807/12/0604–0222\$38.00/0

Accessible online at:
www.karger.com/ann

Prof. Berthold Koletzko
Dr. von Hauner Children's Hospital, University of Munich Medical Centre
Lindwurmstrasse 4
DE–80337 Munich (Germany)
Tel. +49 89 5160 2826, E-Mail office.koletzko@med.uni-muenchen.de

relevant to nutrition trials in this age group for which agreed criteria are needed; (3) agree on an updated 'core data set' that should generally be recorded in nutrition trials in infants and young children, and (4) provide guidance on the use of surrogate markers in paediatric nutrition research. The participants discussed these objectives and agreed to set up six first working groups under the auspices of the Consensus Group on Outcome Measures Made in Paediatric Enteral Nutrition Clinical Trials (COMMENT). Five groups will aim to identify and define criteria for assessing key outcomes, i.e. growth, acute diarrhoea, atopic dermatitis and cows' milk protein allergy, infections and 'gut comfort'. The sixth group will review and update the 'core data set'. The COMMENT Steering Committee will discuss and decide upon a method for reaching consensus which will be used by all working groups and plan to meet again within 2 years and to report and publish their conclusions.

Copyright © 2012 S. Karger AG, Basel

Background

Nutrition during infancy and early childhood (here defined as children aged up to 3 years) must meet the physiological nutrient requirements and support healthy growth and normal development. Considerable evidence has accumulated that infant feeding can markedly influence relevant functional and health outcomes on a short- and long-term basis [1, 2]. Several unresolved questions and controversies exist among paediatricians, scientists and regulatory bodies on the quality and quantity of evidence that is needed to draw firm conclusions on the presence of functional and clinical effects, as well as on the approaches to assess the relevance and impact of such effects for child and population health and for health economic impact. Therefore, a workshop was held in March 2011 under the auspices of the Early Nutrition Academy (www.early-nutrition.org) and the charitable Child Health Foundation (www.kindergesundheit.de), in collaboration with the Committee on Nutrition of the European Society for Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN; www.espghan.org). The main goal of the workshop was to prepare an approach for guidance on the ways to acquire good quality scientific evidence on functional effects of nutritional interventions in infants and young children.

Previously, several position papers and recommendations related to trials evaluating effects of nutritional interventions in this age group have been published [1, 3–8] (table 1).

Despite these guidance documents and numerous published randomised controlled trials (RCTs), conclusions on the efficacy and benefits of many nutritional interventions and innovations used in infancy still remain uncertain and controversial, for example the addition of compounds with proposed probiotic and prebiotic effects to infant foods, as recently highlighted by the ESPGHAN Committee on Nutrition [9].

There is still no agreement within the scientific community about how to best define and measure outcomes used in nutrition trials conducted in infants and young children. For example, a recent systematic review of 138 RCTs reporting on paediatric acute diarrhoea found that 64 different definitions of diarrhoea, 69 definitions of diarrhoea resolution and 46 unique primary outcomes were used [10]. Thus, even in sound clinical trials on paediatric acute diarrhoea, many definitions are heterogeneous, lack evidence of validity and focus on indices that may not be relevant to child health or quality of life. Hence, the conclusions that can be drawn from the accumulated evidence are more restricted than they could otherwise be.

The use of inappropriate outcome measures and/or their definitions may result in misleading information on the relevance of the outcome measure for infant health. It also may result in overestimation, underestimation or failure to reveal potential benefits of the intervention. Moreover, there are unresolved questions and controversies among paediatricians, scientists and regulatory bodies on the quality and quantity of evidence that is needed to draw firm conclusions on the presence of functional effects. The best approaches to assess the relevance and health economic impact of such effects for child and population health are also not clear. For example, questions have been raised on which conditions should be met by clinical trials, whether single or multiple trials are required to draw conclusions and whether documentation of functional effects requires the assessment of clinical endpoints, surrogate markers or the combination of both. These uncertainties could limit available resources for studies on innovation in paediatric nutrition that could benefit child health and well-being.

Considering these issues, the workshop participants decided that consensus on a core set of outcomes with agreed definitions that should be measured and reported in nutrition trials is needed. Since the workshop was arranged under the auspices of the Early Nutrition Academy and ESPGHAN, the group decided to call this initiative the Early Nutrition Academy/ESPGHAN Consensus Group on Outcome Measures Made in Paediatric Enteral Nutrition Clinical Trials (to be known as COMMENT).

Table 1. Position papers and recommendations related to infant nutrition trials

Ref. No.	Title of paper/report	Key conclusions
3	The nutritional and safety assessment of breast milk substitutes and other dietary products for infants: a commentary by the ESPGHAN Committee on Nutrition	modifications of infant feeding regimens and dietary products need to be evaluated in clinical trials meeting accepted standards of scientific methodology
7	Infant Formula: Evaluating the Safety of New Ingredients	existing guidelines and regulations for evaluating the safety of conventional food ingredients (e.g. vitamins and minerals) added to infant formulas are not sufficient to address the diversity of potential new ingredients proposed by manufacturers to develop formulas that mimic the perceived and potential benefits of human milk
4	Characterisation of infant food modifications in the European Union	general concepts on evaluating innovations and on establishing evidence for benefits collated by a multidisciplinary expert group
5	Core data for nutrition trials in infants: a discussion document – a commentary by the ESPGHAN Committee on Nutrition	proposed, for general discussion, a list of core data that should be recorded in nutrition trials and suggested principles for the identification of data required for studies with different outcome measures
6	Report of the Scientific Committee on Food on the Revision of Essential Requirements of Infant Formulae and Follow-on Formulae	guidance on the evaluation of compositional modifications of infant formulae or follow-on formulae; general agreement that the choice of measures of efficacy and safety will be determined by the nature of the study and the hypothesis being tested; examples for growth and neurodevelopment were provided

The COMMENT Approach

Problems to Be Addressed

The workshop participants agreed that problems related to the choice and definition of outcomes assessing the effects of the addition of new ingredients to infant formulae need to be urgently addressed. The recent comment of the ESPGHAN Committee on Nutrition on supplementation of infant formulae with compounds considered to have pre- and probiotic effects demonstrates these problems [9]. Growth is typically assessed by anthropometric measurements, namely weight, length and head circumference. However, interpreting studies on the effects of probiotic or prebiotic supplementation of infant formulae on growth is difficult for several reasons. Firstly, very few studies have analysed the effects of any specific probiotic strain and/or prebiotic compound. Secondly, several studies included a sample size that was too small, with insufficient power to identify relevant ef-

fects on growth as well as other outcomes. Finally, the follow-up periods in the trials were often too short. The choice and definitions of clinical outcomes differed, even those relating to the same domain (e.g. gastrointestinal infections). Even if the same outcomes were measured, the chosen criteria for these outcomes were heterogeneous, often not widely agreed upon or just not reported.

Extrapolation of these specific conclusions relating to probiotics or prebiotics to other modifications of nutritional interventions in infants and young children, e.g. the addition of any innovative ingredient to infant formula, allowed agreement to be reached on several issues. The choice of outcomes, definitions of outcomes and instruments employed in trials to evaluate the safety and efficacy of infant formulae supplemented with innovative ingredients are heterogeneous, often lack evidence of reliability and validity and sometimes focus on indices that may not be relevant to infants and their caregivers [9].

Mindful of this background, the COMMENT workshop participants agreed on the objectives outlined below.

Overall Objectives of COMMENT

- (1) To provide guidance on the level of quality and quantity of evidence needed to justify conclusions of effects of nutrition in infants and young children on biological function and clinical endpoints;
- (2) to agree on a range of outcome measures relevant to nutrition trials in infants and young children for which agreed criteria are needed;
- (3) to agree on an updated 'core data set' that should generally be recorded in nutrition trials in infants and young children, and
- (4) to provide guidance on the use, strengths and limitations of surrogate markers.

The Kick-Off Workshop in Tutzing 2011

The workshop spanned 2.5 days in March 2011 and was held in Tutzing near Munich, Germany. Workshop participants included researchers with knowledge in different areas including clinical trials, paediatric allergology, infectious diseases, gut physiology, neonatology, neurodevelopment, nutrition and metabolism, paediatric gastroenterology and regulatory evaluation. Researchers working in the infant food industry and in a company developing and evaluating biomarkers were invited to attend on the first day only to provide presentations on their experience and to answer questions regarding pre-clinical and clinical evaluation. The group discussions and conclusions were held on the second and third day, without the presence of any representatives of industry. A list of the attendees from industry and their affiliations can be found at the end of this paper.

The workshop participants reviewed and discussed the impressive achievements of the Outcome Measures in Arthritis Clinical Trials group, an international, informally organized network initiated in 1992 to improve outcome measurement in clinical studies in rheumatology (including rheumatoid arthritis, osteoarthritis, osteoporosis, ankylosing spondylitis and gout). Chaired by an executive committee, it organises consensus conferences in a 2-yearly cycle. Data-driven recommendations are prepared and updated by expert working groups [11, 12]. The workshop participants agreed that COMMENT should try to set similar objectives and attempt to function in a similar way for investigations in the field of early nutrition.

Discussion on Objective 1: To Provide Guidance on the Level of Quality and Quantity of Evidence Needed to Justify Conclusions on Functional and Clinical Effects of Nutrition in Infants and Young Children

What guidance can be given for conducting good quality studies in infant nutrition to demonstrate functional effects?

Type of Study and Type of Evidence

The workshop participants agreed that the preferred study design depended primarily, and most importantly, on the question which was being addressed. For example, although double-blind RCTs are generally accepted as the 'gold standard', some research questions (such as the benefits of breastfeeding) cannot always be answered by a randomised study, and other types of evidence (e.g. observational studies of high quality with adequate adjustment for confounding) must be considered. Moreover, not all nutritional interventions can be securely double masked (e.g. protein hydrolysates may differ in taste and induced stool characteristics, and prebiotics or dietary fibre may induce softer stools). The workshop participants concluded that obtaining a full body of good methodological quality evidence from all sources, with consistent results, is most important. However, a larger quantity of poor evidence can never replace a smaller quantity of good quality evidence.

Understanding the mechanisms of beneficial effects is important, and this can help in the design of better RCTs. However, there are many examples of effects for which the detailed mechanisms are unknown (e.g. many of those in the cancer field). Therefore, showing a plausible mechanism or measuring a biomarker reflecting underlying mechanisms should be considered a very useful addition to good quality trial data but not a prerequisite for demonstrating effects of a nutritional intervention.

The workshop participants agreed there was no conclusive current information that post-marketing surveillance data (as obtained, say, by the infant food industry after a product launch) were valuable. So far, there is no indication that post-marketing surveillance data are useful for documenting the benefit or safety of food products for infants and young children. Agreed standards have never been set for the conduct or reporting of post-marketing surveillance, and there are too many confounding factors, such as the education level of subjects.

Trial Planning, Execution and Reporting

The workshop participants agreed that a consistent approach to the planning, conduct and reporting of RCTs is needed, not only with regard to study design but also in terms of the consistency of choosing study and control populations, taking samples from them and the storage of those samples.

Trial and protocol registration prior to subject recruitment is now an essential requirement for any trial in human subjects to avoid unintended duplication of trials. It prevents the suppression and selective reporting of results based on the interests of funders, unacknowledged alterations of pre-specified outcome measures and the failure to report relevant adverse events [13]. The indispensable need for such registration is also supported by the conclusions of a recent observational study of safety and efficacy trials for five drug categories. This showed that only two thirds of the results had been published within 2 years of the completion of the registered drug trials. Those funded by industry were less likely to be published, and those published were more likely to report positive outcomes than were trials funded by other sources [14].

The Consolidated Standards Of Reporting Trials statement gives guidelines for reporting trials [15]. These guidelines were developed for drug trials, but several papers have provided guidance on designing clinical trials for specific functional food ingredients [16–18]. Consensus guidelines have recently been collated for designing, conducting and reporting human intervention studies to evaluate the health benefits of foods [19]. These can be followed, to a great extent, in trials conducted with infants and young children.

Such consistency in all these aspects of trial design will not only ensure that high-quality trials are produced but will also ensure that banked samples from fully documented subjects can be available for future studies.

Study, Control and Reference Populations

Matched Control Population

Considering the study population, the workshop participants discussed whether it needed to be identical to the target population. They agreed that, in some cases, it may be possible to perform studies in a different age group. For example, results from comparative studies on the bioavailability of nutrients from different food sources and on their safety may often be suitable for extrapolation to other age groups. However, if studies are performed in a population with a high probability for an outcome (e.g. family history of allergy, family history or

genetic markers of celiac disease or type 1 diabetes, population with a high diarrhoea prevalence), care must be taken if the results are extrapolated to the general population of healthy infants and young children.

The gold-standard RCT requires a perfectly matched and randomised control population, but the workshop participants agreed that, in a few cases, some flexibility with respect to the choice of reference data could be allowed. For example, in growth studies, the growth of the study population could be compared with growth curves from standard populations, including the recently published WHO growth standards [20]. However, the legitimacy of this strategy would depend on the research question being asked. For example, evaluation of the suitability of a particular novel ingredient would generally require an RCT with a control population fed the same formula but without the reference ingredient, whereas a formula modification aiming at matching the growth rate of breastfed populations might be compared to WHO growth standards.

The workshop participants agreed that there are scientific and ethical considerations that need to be taken into account in comparison of populations with different ethnic cultures and lifestyles, such as those from developing countries and less privileged segments of the population. Different genetic and lifestyle backgrounds may mean that outcomes such as growth curves could not always be extrapolated to studies in Europe.

What is ‘standard feeding’ of the control group? For example, in allergy prevention trials with modified formula in children with a documented family history of allergic disease, is standard feeding a regular cows’ milk formula or a formula based on protein hydrolysates with previously documented allergy risk reduction? The answer may differ depending on the standard feeding choices in various settings or countries. However, it was agreed that ideally the control and study interventions should only differ in one component, namely the functional component of interest.

Reference Population

Another alternative to the matched control population is the use of a reference population. Are populations of breastfed infants considered the ‘gold standard’, and should they generally be used as the reference? Should all studies aim at inclusion of reference breastfed populations? Although the workshop participants agreed that reference breastfed populations are often useful, they recognised the difficulties posed since consistently there are large sociodemographic differences relevant to health

between populations of breastfed and non-breastfed infants. The workshop participants agreed that the choice of whether or not to include a breastfed group depends on the research question. If safety or efficacy of an added component is to be tested, it may suffice to compare two formula groups fed formula with or without the added component. If, however, the goal is to demonstrate similarity to clinical outcomes in breastfed populations, then inclusion of a reference breastfed group should be considered. In some instances, comparison of results to data from previously studied reference breastfed populations or growth reference data of breastfed populations can be adequate.

Other Considerations

Practicalities might also determine the starting age for trial participants. Clinical studies that assess parameters such as growth, and safety and efficacy of an infant formula may provide the most informative results when the formula is fed as the sole source of nutrition from birth. This is the period with the most rapid growth, highest nutrient requirements per kilogram of body weight and highest sensitivity to an unbalanced supply. However, in Europe, it has now become very difficult to enroll infants fed formula from birth since, fortunately, most healthy infants are breastfed after birth. The established approach for an infant growth study is to power it to detect differences in growth using 0.5 standard deviation in a study starting from birth with a duration of 3 months [5]. If enrolment into a study has to be started at a later time point after an initial period of breastfeeding, e.g. during the second or third month of life, studies may need to be powered to detect smaller effect sizes to document potential effects of food components on growth at an age with lower nutrient needs than those immediately after birth.

As there is no scientific rationale to support a strict distinction between development in infants under and above 6 months of age with respect to the effect of factors such as nutrients on growth and development, to what extent can data collected in infants 0–6 months of age be extrapolated to infants 6–12 months of age? The workshop participants agreed that this should be decided on a case-by-case basis and that it was best left for discussion in the specific working groups (see below).

The importance of deciding upon a statistical plan for analysing results at the outset of a trial was emphasised. The level for statistical significance should also be agreed at this point. The workshop participants agreed that this level did not always have to be $p < 0.05$, especially in

studies of healthy populations, but other levels of significance might be chosen depending on the given conditions.

Transparency of approach, namely registration and/or publication of the study design and primary outcome measures at the outset, is essential. There is no problem with assessing more than one health outcome in a trial, as long as this is decided upon and transparently reported at the outset. Assessing multiple primary outcomes should be reflected in the power analysis and increases the number of subjects needed in the trial. Analysing secondary outcomes is also permissible as long as this is made clear at the outset.

Open-Label Extension to Trials

Quite often subjects participating in RCTs are followed up further, after the un-blinding of the trial, to explore potential longer-term effects of the intervention. Such ‘open-label’ extensions of RCTs are obviously subject to potential bias due to lack of blinding. An even more important threat to validity may be dropouts from the sample randomised in the preceding RCT [21, 22]. Nonetheless, follow-up of the long-term health effects of early nutritional interventions is encouraged, since these outcomes are of great public health interest and this type of information is currently still scarce.

Replication of Trials

The workshop participants discussed whether replication of clinical trial results in an independent second trial was generally necessary. Replication may not be easy if there is lack of scientific novelty or little interest from potential sponsors if a commercially available ingredient showed positive effects in a previous study. However, its importance is demonstrated by contrasting results in replicated studies. For example, Kopp et al. [23] evaluated the effect of using *Lactobacillus* GG for the primary prevention of atopic dermatitis following a protocol almost identical to that used in a study [24] which had reported a preventive effect on the development of atopic eczema. However, the later investigators found no benefit of *Lactobacillus* GG use in the prevention of atopic dermatitis [23].

Consequently, the workshop participants agreed that sometimes one trial may suffice, but sometimes more than one trial may be necessary. At least two independent adequate trials with similar outcomes are often needed to provide convincing evidence and to change practice, but this is not always achievable, particularly if a large sample size or long-term follow-up is needed to test a specific hy-

pothesis. The participants discussed whether replication in a second, more or less identical, study and burdening another population with, say, blood samples, was ethical? It was agreed that the best option was to repeat the trial but to modify it so that, say, another primary outcome is added to those tested before.

Growth Studies

Growth has always been an important outcome in nutritional intervention studies. Traditionally, it has been regarded as a safety outcome to ensure that new nutritional products adequately supported growth during the important first years of life. With the obesity epidemic and the increasing focus on long-term effects of early nutrition and early growth, interpretation of growth data during the first years of life has become more complex. The workshop discussion suggested that perhaps growth studies are not always necessary when functional effects are being investigated, and that perhaps growth studies are only needed under specific circumstances, for example the modification of protein content or quality, or the addition of complex carbohydrates, which might modify the bioavailability of other formula compounds.

Safety

The workshop participants stressed the overall importance of checking safety at the same time as functionality.

Comparison between Quality and Quantity of Evidence to Draw Conclusions in the Pharma Industry and the Infant Food Industry

The participation of infant food industry representatives on day 1 allowed the discussion of the following question: 'Is the same high evidence level necessary to document functional effects of food, e.g. the promotion of a functional development or of the risk reduction of an allergic manifestation, as compared to the effects of therapeutic drugs?'

The workshop participants agreed that the food industry is not the same as the pharma industry but that the 'controls/guidance' for the *infant* food industry have many parallels with the pharma industry [25]. This is due to the sensitive age group and the need to ensure the highest possible level of safety for feeding infants and young children. It is therefore useful to consider pharma, with its tight regulations, as a model. Having said this, the workshop participants recognised that the context within which the infant food industry operates is much more

complicated than that in which the pharma industry operates for the following reasons:

- A reductionist approach is usually used for pharma, but a holistic approach is always needed for food. However, investigating the effects of feeding infants and follow-up formulas is simpler than looking at the effects of food on adults because the diet is less heterogeneous.
- Pharma products are generally used for the treatment of disease in those who are already ill, whereas functional food products (including those for infants) are mainly targeted at healthy people, aiming at securing optimal health and development and lowering disease risk.
- Pharma products lend themselves more to being tested in double-blind RCTs compared with food ingredients, which are sometimes difficult to mask. Compliance tends to be better in drug trials, in which participating subjects suffer from a burden of disease and are usually closely supervised by a health care professional.
- The influence of age and ethnicity is often less important in pharma trials compared with their influence in food trials in infants and young children, where age and ethnicity are often closely associated with nutrient supply and status.

Within the infant food industry, demonstration of safety has always been paramount, but the current climate demands proof of efficacy much more than ever before.

Discussion on Objective 2: To Agree on a Range of Outcome Measures Relevant to Nutrition Trials in Infants and Young Children for Which Agreed Criteria Are Needed

Which outcomes could be assessed in clinical trials?

The presentations and discussions covered a range of outcome measures which could be classified as shown in table 2.

Points of Ambiguity and Questions about Outcome Measures

The workshop participants recognised that there is confusion and ambiguity between clinical endpoints, intermediary endpoints and markers. They agreed that the following are all relevant questions: Which clinical outcome parameters can be considered indicative for improving function and subsequently beneficial for infant

health? Can we provide criteria for effect sizes that are relevant? Should we consider the number needed to treat? Which factors in infancy/early childhood should be considered to predict later disease risk? What is the value of predefined secondary outcomes compared to predefined primary outcomes?

General Principles for Outcome Measures

The workshop participants decided that it would be best to discuss these questions in detail in specific working groups devoted to particular outcomes (see below). However, they were able to agree some general principles:

- it is a priority to agree upon common definitions for outcome measures, and this can realistically be achieved in parallel working groups;
- a good outcome measure which is easy to measure and is clearly defined (diarrhoea would make a good example if this can be achieved) will make the use of markers redundant for that particular health outcome, and
- markers for health outcomes with good outcome measures could still be needed if we need to use that marker to predict future events (like growth).

Discussion on Objective 3: To Agree on an Updated 'Core Data Set' That Should Generally Be Recorded in Nutrition Trials in Infants and Young Children

A discussion document from the ESPGHAN Committee on Nutrition published in 2003 has suggested a core data set for nutrition trials in infants [5]. This document suggests checklists for identification data and core data for all studies of infants, as well as for studies of infants born preterm [5]. It has been used and referred to by several investigators.

It was agreed that the COMMENT working group will join forces with the ESPGHAN Committee on Nutrition in order to publish a revised ESPGHAN position paper on these matters.

An updated core data set will ensure that full details of body composition with a full clinical chemical profile are always collected in trials in infants and young children. This will then facilitate comparison of results from studies, for example in meta-analyses. However, all workshop participants agreed that, although it was essential to collect comprehensive core data, in doing so, the burden on parents and children must be kept to a minimum.

Table 2. Classification of outcome measures

<i>Growth</i>	Functional outcomes, development and health, e.g. gut physiology, permeability or inflammation, stool bacteria, stool patterns
<i>Neurodevelopment</i>	<ul style="list-style-type: none"> – Sensory function (e.g. vision, auditory) – Motor function – Cognitive function
<i>Immune response</i>	<ul style="list-style-type: none"> – Clinical endpoints versus biomarkers
<i>Disease risk reduction</i>	<ul style="list-style-type: none"> – Infectious diseases, e.g. gastroenteritis, upper respiratory tract infection, urinary tract infection – Eczema/atopic eczema – Food allergy – Other allergic disorders, e.g. asthma, wheezing, eosinophilic gastroenteropathy/oesophagitis – Coeliac disease – Type 1 diabetes – Autoimmune diseases, inflammatory bowel diseases – Prematurity-related disorders (necrotising enterocolitis, intestinal failure) – Other chronic diseases (obesity, cardiovascular disease)

Discussion on Objective 4: To Provide Guidance on the Use, Strengths and Limitations of Markers

Background

The workshop participants reviewed how the nature and characterisation of markers, in the context of exploring the evidence needed to support health claims on foods, were initially addressed initially in the EC project Functional Food Science in Europe [26] and then in the EC project Process for the Assessment of Scientific Support for Claims on Foods [27], which explored further how the markers would be used as part of the strategic essence of the portfolio of evidence submitted to substantiate a health claim.

The two concerted actions deliberately avoided using the term 'biomarker' because they wish to convey the message that any relevant phenomenon or outcome could serve as a 'marker'. Markers can be used to give an indication of intake (external exposure or dosage), body burden (internal exposure, 'status'), systemic intermediate metabolism and the production of metabolites (both active and inactive) or intermediate functional effects, leading to the ultimate desired outcome. If the latter cannot be measured, then a marker can be used as a 'surrogate endpoint'.

This sequence is completely analogous to the absorption, distribution, metabolism and excretion model of pharmacokinetics or toxicokinetics, and strategies for identifying, if necessary, intermediate markers for nutritional studies can be learned from experience in these fields.

If markers represent an event directly or causally involved in the process under study, they have been called 'factors', whereas if a direct functional correlation has not been demonstrated and they may in fact be bystanders or associated phenomena, they are called 'indicators'. Discussions about whether or not markers are indicators or factors come from discussions about observational epidemiological studies and are not so relevant to discussion of trials.

Thus, there is a need to give serious thought to the validity of markers in relation to their expected purpose in any experimental study. The fundamental criteria for markers in any clinical study are generic. Markers can be chemical, biochemical, physiological (i.e. addressing performance or adaptive capacity), observational or behavioural, in short, anything that can be objectively measured. As such, they therefore should be feasible, valid, reproducible, sensitive and specific. The following characteristics apply equally well to all types of marker:

- markers should be measurable at relatively immediate time points to be useful in assessing interventions in a reasonable timescale;
- markers should be rigorously validated and amenable to standard quality control procedures;
- they should be clearly linked to the phenomena involved in the biological process being studied; ideally, usually as part of the validation, markers other than the endpoint should have an obvious role in the mechanistic pathway leading to the outcome being studied;
- markers should usually be gained from an easily accessible material, by a minimally invasive and ethical procedure;
- markers may need to be validated from mechanistic studies in animal models and cellular studies;
- markers need not be single static measurements; they may represent dynamic responses such as clearance studies, changes in enzyme function or signal transduction, and
- static and dynamic markers might also be based on objective measurements of psychological and physical performance and subjective assessments of quality-of-life or similar outcomes.

Above all, investigators should focus on study design and on the selection of appropriate phenomena or characteristics that can be measured to investigate or demon-

strate a specific causal relationship. Thus, markers and their roles can be identified within the context of the study design.

Workshop Decisions

The workshop participants decided they could only give some general directions on markers at this stage. Firstly, the workshop participants agreed that they would use the term 'marker' rather than 'biomarker' for the reasons given above.

Secondly, it is important to distinguish markers of exposure (external and internal) to a functional component from markers which could be predictive of the effect of that component on functional outcome or reduced risk of disease. A variety of exposure markers for dietary intake of nutrients and non-nutrients have been reported and to some extent validated for infants and young children [28–30].

The main point is to choose markers that have been proven to correlate with clinical outcome. However, there are far fewer data on markers for outcome in early childhood than there are for adults (for example, in adults, high serum low-density lipoprotein cholesterol levels are accepted as a predictor of cardiovascular disease risk). At present, there are only limited validated markers of clinical outcomes relevant for infant nutrition studies (e.g. growth, adiposity, allergy). Whenever markers exist, based on known mechanisms of a dietary component effect on a clinical endpoint, it would obviously be most useful to include these in any clinical trials. Since this is not often the case, the current focus is on the assessment of clinical endpoints. Clearly, identification of a marker is not a prerequisite for conclusively demonstrating efficacy based on documentation of clinical endpoint effects.

New exciting opportunities for developing informative markers may arise from systems biology (metabolomic and bioinformatic methodology) and mathematic modelling, which might help strategically develop metabolic maps and thus identify good exposure and predictive markers.

Next Steps

Working Groups

The workshop participants agreed that six working groups should be set up as soon as possible and that others (such as immune function) might be considered later. Leaders of these working groups were agreed and are shown in table 3.

The working group leaders will recruit other experts as needed, but the working groups will remain small, with usually no more than 4–5 members. Their aim is to identify and define criteria for assessing key outcomes. The groups will perform a systematic review at the outset to inform on the current heterogeneity of outcomes. The COMMENT Steering Committee will discuss and decide upon a method for reaching consensus which will be used by all working groups. A section of the ESPGHAN website will be dedicated to COMMENT to enable good communication between working groups and their members.

Consensus-Forming Techniques

The COMMENT Steering Committee will discuss and decide upon a method for reaching consensus which will be used by all working groups. Several were suggested at the workshop, including the nominal group technique, the Delphi technique and the decision support matrix technique. The working groups and the Steering Committee will decide on the strategies for future meetings and publications.

Dissemination and Exploitation of COMMENT Discussions and Consensus to Stakeholders

The COMMENT workshop participants agreed that their conclusions should reach clinicians/scientists who are involved in preclinical and clinical studies, scientific societies and bodies providing recommendations (e.g. ESPGHAN, American Academy of Pediatrics), risk assessment bodies (e.g. European Food Safety Authority, national bodies such as Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (France), Bundesinstitut für Risikobewertung (Germany), Food and Drug Administration (USA), Food Safety Authority (UK)) and risk management bodies (e.g. European Commission, Codex Alimentarius, national bodies).

References

- Koletzko B, Aggett PJ, Bindels JG, Bung P, Ferre P, Gil A, et al: Growth, development and differentiation: a functional food science approach. *Br J Nutr* 1998;80(suppl 1): S5–S45.
- Koletzko B, Decsi T, Molnar D, de la Hunty A (eds): Early Nutrition Programming and Health Outcomes in Later Life: Obesity and Beyond. *Advances in Experimental Medicine and Biology*. Springer, 2009, vol 646:1–196.
- Aggett PJ, Agostini C, Goulet O, Hernell O, Koletzko B, Lafeber HL, Michaelsen KF, Rigo J, Weaver LT; European Society of Pediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) Committee on Nutrition: The nutritional and safety assessment of breast milk substitutes and other dietary products for infants: a commentary by the ESPGHAN Committee on Nutrition. *J Pediatr Gastroenterol Nutr* 2001;32:256–258.
- Koletzko B, Ashwell M, Beck B, Bronner A, Mathioudakis B: Characterisation of infant food modifications in the European Union. *Ann Nutr Metab* 2002;46:231–242.
- Aggett P, Agostini C, Axelsson I, Goulet O, Hernell O, Koletzko B, Lafeber HN, Michaelsen KF, Morley R, Rigo J, Szajewska H, Weaver LT; ESPGHAN Committee on Nutrition: Core data for nutrition trials in infants: a discussion document – a commentary by the ESPGHAN Committee on Nutrition. *J Pediatr Gastroenterol Nutr* 2003;36:338–342.

Table 3. Working groups that were agreed to be set up

Working group	Proposed leader
Growth	Prof. Kim Fleischer Michaelsen
Acute diarrhoea	Prof. Hania Szajewska
Atopic dermatitis and cows' milk protein allergy	Prof. Christophe Dupont
Infections (e.g. upper/lower respiratory tract, urinary tract)	Prof. Alfredo Guarino
'Gut comfort' (e.g. colic, constipation, bloating)	Prof. Marc Beningna
Review of core data set	Prof. Bert Koletzko

Disclosure Statement

The workshop was supported by unrestricted educational grants to the Early Nutrition Academy from Abbott Nutrition, Danone Research, Mead Johnson Nutrition and Nestlé Nutrition. Partial financial support from the Commission of the European Communities (Research and Technological Development Programme 'Quality of Life and Management of Living Resources', within the 7th Framework Programme, Project EarlyNutrition, FP7-289346) for the writing of this report is gratefully acknowledged. This article does not necessarily reflect the views of the Commission and in no way anticipates the future policy in this area.

The following industry observers attended the workshop for one day only: Martine Alles, PhD, Danone Research; Leslie Curry, PhD, Abbott Nutrition; Marie-Odile Galing, Nestlé Nutrition; Johan Garssen, PhD, Danone Research; Udo Herz, PhD, Mead Johnson; Dineke Klaassen, MSc, Danone Research; Mike Possner, MD, Nestlé Nutrition; Ricardo Rueda, PhD, Abbott Nutrition; Peter van Dael, PhD, Mead Johnson, and Klaus Weinberger, PhD, Biocrates Life Sciences AG.

- 6 Scientific Committee on Food: Report of the Scientific Committee on Food on the Revision of Essential Requirements of Infant Formulae and Follow-on Formulae. Brussels, Scientific Committee on Food, 2003.
- 7 Committee on the Evaluation of the Addition of Ingredients New to Infant Formula: Infant Formula: Evaluating the Safety of New Ingredients. Washington, National Academies Press, 2001.
- 8 Aggett PJ, Agostoni C, Axelsson I, Edwards CA, Goulet O, Hernell O, Koletzko B, Lafeber HN, Micheli JL, Michaelsen KF, Rigo J, Szajewska H, Weaver LT; ESPGHAN Committee on Nutrition: Nondigestible carbohydrates in the diets of infants and young children: a commentary by the ESPGHAN Committee on Nutrition. *J Pediatr Gastroenterol Nutr* 2003;36:329–337.
- 9 Braegger C, Chmielewska A, Decsi T, Kolacek S, Mihatsch W, Moreno L, Pieścik M, Puntis J, Shamir R, Szajewska H, Turck D, van Goudoever J; ESPGHAN Committee on Nutrition: Supplementation of infant formula with probiotics and/or prebiotics: a systematic review and comment by the ESPGHAN committee on nutrition. *J Pediatr Gastroenterol Nutr* 2011;52:238–250.
- 10 Johnston BC, Shamseer L, da Costa BR, Tsuyuki RT, Vohra S: Measurement issues in trials of pediatric acute diarrheal diseases: a systematic review. *Pediatrics* 2010;126:e222–e231.
- 11 Brooks P, Boers M, Simon LS, Strand V, Tugwell P: Outcome measures in rheumatoid arthritis: the OMERACT process. *Expert Rev Clin Immunol* 2007;3:271–275.
- 12 Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L: OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* 2007;8:38.
- 13 Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC: The ClinicalTrials.gov results database – update and key issues. *N Engl J Med* 2011;364:852–860.
- 14 Bourgeois FT, Murthy S, Mandl KD: Outcome reporting among drug trials registered in ClinicalTrials.gov. *Ann Intern Med* 2010;153:158–166.
- 15 Schulz KF, Altman DG, Moher D: CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332.
- 16 AbuMweis SS, Jew S, Jones PJ: Optimizing clinical trial design for assessing the efficacy of functional foods. *Nutr Rev* 2010;68:485–499.
- 17 Klein MA, Nahin RL, Messina MJ, Rader JL, Thompson LU, Badger TM, et al: Guidance from an NIH workshop on designing, implementing, and reporting clinical studies of soy interventions. *J Nutr* 2010;140:1192S–1204S.
- 18 Shane AL, Cabana MD, Vidry S, Merenstein D, Hummelen R, Ellis CL, et al: Guide to designing, conducting, publishing and communicating results of clinical studies involving probiotic applications in human participants. *Gut Microbes* 2010;1:243–253.
- 19 Welch RW, Antoine JM, Berta JL, Bub A, de Vries J, Guarner F, Hasselwander O, Hendriks H, Jäkel M, Koletzko BV, Patterson CC, Richelle M, Skarp M, Theis S, Vidry S, Woodside JV; International Life Sciences Institute Europe Functional Foods Task Force: Guidelines for the design, conduct and reporting of human intervention studies to evaluate the health benefits of foods. *Br J Nutr* 2011;106(suppl 2):S3–S15.
- 20 de Onis M, Siyam A, Borghi E, Onyango AW, Piwoz E, Garza C: Comparison of the World Health Organization growth velocity standards with existing US reference data. *Pediatrics* 2011;128:e18–e26.
- 21 Megan B, Pickering RM, Weatherall M: Design, objectives, execution and reporting of published open-label extension studies. *J Eval Clin Pract* 2012;18:209–215.
- 22 Fewtrell MS, Kennedy K, Singhal A, Martin RM, Ness A, Hadders-Algra M, et al: How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? *Arch Dis Child* 2008;93:458–461.
- 23 Kopp MV, Hennemuth I, Heinzmann A, Urbanek R: Randomized, double-blind, placebo-controlled trial of probiotics for primary prevention: no clinical effects of Lactobacillus GG supplementation. *Pediatrics* 2008;121:e850–e856.
- 24 Kalliomaki M, Salminen S, Arvilommi H, Kero P, Koskinen P, Isolauri E: Probiotics in primary prevention of atopic disease: a randomised placebo-controlled trial. *Lancet* 2001;357:1076–1079.
- 25 Georgiou NA, Garssen J, Witkamp RF: Pharma-nutrition interface: the gap is narrowing. *Eur J Pharmacol* 2010;651:1–8.
- 26 Diplock A, Aggett P, Ashwell M, Bornet F, Fern E, Roberfroid M: Scientific concepts of functional foods in Europe: consensus document. *Br J Nutr* 1999;81:S1–S27.
- 27 Aggett PJ, Antoine JM, Asp NG, Bellisle F, Contor L, Cummings JH, Howlett J, Müller DJ, Persin C, Pijls LT, Rechkemmer G, Tuijtelars S, Verhagen H: PASSCLAIM: consensus on criteria. *Eur J Nutr* 2005;44(suppl 1):i5–i30.
- 28 Fekete K, Marosvolgyi T, Jakobik V, Decsi T: Methods of assessment of n-3 long-chain polyunsaturated fatty acid status in humans: a systematic review. *Am J Clin Nutr* 2009;89:2070S–2084S.
- 29 Hooper L, Ashton K, Harvey LJ, Decsi T, Fairweather-Tait SJ: Assessing potential biomarkers of micronutrient status by using a systematic review methodology: methods. *Am J Clin Nutr* 2009;89:1953S–1959S.
- 30 Lowe NM, Fekete K, Decsi T: Methods of assessment of zinc status in humans: a systematic review. *Am J Clin Nutr* 2009;89:2040S–2051S.