Nutrition Bulletin

DOI: 10.1111/nbu.12241

Guidance for the design and implementation of human dietary intervention studies for health claim submissions

A. Lucey, C. Heneghan and M. E. Kiely

Cork Centre for Vitamin D and Nutrition Research, School of Food and Nutritional Sciences, University College Cork, Cork, Republic of Ireland

Abstract

The recently completed EU-funded *BACCHUS* Integrated Project (grant number: 312090) has focused on identification of cause-and-effect relationships between consumption of bioactive peptides and polyphenols and physiological effects on cardiovascular health. An important *BACCHUS* output is a toolkit to support the needs of small- and medium-sized enterprises (SMEs) in the food sector considering making health claims on their products. The toolkit draws together best practice guides, a bioactives database, an intake assessment tool and an e-learning platform. This paper focuses on the *Best Practice Guide for Human Dietary Intervention Studies*. The toolkit has been developed with SMEs in mind but this guidance is likely to have value far beyond the needs of businesses, in particular being a useful resource for students, early-career scientists and others new to the design and implementation of dietary intervention studies. The aim of this article is to share the principles of the guide with a broader audience.

Keywords: dietary intervention studies, food bioactives, health claims, randomised controlled trials, study design

Introduction

A Regulation on Nutrition and Health Claims [Regulation (EC) No. 1924/2006] (European Commission 2007) came into force in the European Union (EU) in 2007 (Buttriss 2015). Health claims on food products aim to inform the public about the health benefits of the product. The Regulation is designed to protect consumers against misleading claims, by ensuring that nutrition and health claims used in Europe are scientifically valid, harmonise claims made across the EU and encourage innovation in the food industry.

The Regulation covers all foods, drinks and dietary supplements sold in the EU. It applies to all commercial communications including the food label itself, advertorials and other promotional materials. The claim must apply to the product as consumed, prepared according to the manufacturer's instructions and the effects described in the claim must be understandable to consumers.

The European Food Safety Authority (EFSA), an independent scientific body whose role is to provide scientific advice to the European Commission (EC), carries out the scientific assessment of evidence for health claims. The role of EFSA in risk assessment is separate from the role of the EC in risk management. In the context of health claims, this means that, at the request of the EC, EFSA provides a scientific opinion

Correspondence: Professor Mairead Kiely, Professor of Human Nutrition, Cork Centre for Vitamin D and Nutrition Research, School of Food and Nutritional Sciences, Maternal and Child Nutrition Program Lead, Irish Centre for Fetal and Neonatal Translational Research, University College Cork, Cork, Republic of Ireland.
E-mail: m.kiely@ucc.ie

on the evidence in support of a specific claim, which the EC then uses to decide whether to accept or reject a claim, alongside consideration of anticipated consumer understanding of the claim.

The 4-year *BACCHUS* project (grant number: 312090) was funded by the EU to develop tools and resources that will facilitate the generation of robust and exploitable scientific evidence that can be used to support claims of a ***cause-and-effect*** (note that text in bold and italics is defined in a glossary at the end of the paper) relationship between consumption of bioactive peptides and polyphenols and beneficial physiological effects related to cardiovascular health in humans. A focal point of the project is a toolkit that has been developed to meet the needs of food-based small- and medium-sized enterprises (SMEs), especially those seeking practical guidance on development of successful health claim dossiers (www.bacchus-fp7.eu/Resources/). Online surveys and one-to-one discussions were utilised to evaluate the toolkit throughout its development to ensure that it meets the needs of the end user. The individual components of the toolkit includes a *Best Practice Guide for Health Claims* (production of which was led by the British Nutrition Foundation), presented in two parts. Part I contains information that relates to the topic of health claims in general and part II covers areas that specifically relate to the *BACCHUS* project [*i.e.* claims related to polyphenols, bioactive peptides and cardiovascular disease (CVD)]. The Guide focuses on official regulations and guidance documents from the EC and EFSA, as well as looking at the outcomes of assessments of relevant claims applications by EFSA and subsequent approval or rejection by the EC. It also contains a health claims template that outlines the information needed for a dossier, which is intended to provide a first step in collating the information required to prepare a health claim dossier.

Also included in the toolkit are the *eBASIS* bioactives database (developed by Institute of Food Research, University College Cork and *EuroFIR*), the *eBASIS-Creme* intake assessment tool (developed by Creme Global and *EuroFIR*), the *Best Practice Guide for Human Dietary Intervention Studies* (developed by University College Cork and the main subject of this paper) and an e-Learning platform (developed by Wagralim). The toolkit components have been brought together on the *BACCHUS* website, with the resources available via registration (www.bacchus-fp7.eu). When the project ends, the resources are to be maintained publicly beyond the lifespan of the project through a dedicated microsite that will be maintained and hosted by *EuroFIR* (www.eurofir.org).

Pertinent studies in humans are an absolute requirement for the scientific substantiation of health claims, and pertinent human ***efficacy studies*** are at the top of the hierarchy that informs decisions on substantiation (EFSA 2016). Dietary intervention studies must be well designed in order to optimise the quality of the data they provide. A key component of the EFSA health claim evaluation process is the provision of robust evidence clearly demonstrating a physiological cause-and-effect relationship, which is clinically relevant to human health from at least one well-designed dietary intervention study or randomised controlled trial (RCT). In addition to the food or ***food constituent*** under review being comprehensively characterised, all data must be sufficiently detailed to identify the strength, consistency, specificity, dose–response and biological plausibility of the proposed relationship between food and health.

The *Best Practice Guide for Human Dietary Intervention Studies* features a comprehensive set of guidelines that consider the major design elements of dietary interventions with a view towards identifying optimal approaches, as well as identifying gaps and challenges in developing protocols for dietary intervention studies. This user-friendly resource was developed for SMEs to support the submission of a health claim for foods/food-derived bioactives related to cardiovascular health. Guidance is provided at an introductory but comprehensive level and is based on authoritative sources, as well as the experience of the authors who led the human studies work package within the *BACCHUS* project. *BACCHUS* scientists undertook six randomised placebo-controlled dietary intervention studies investigating the effects of fruit-derived polyphenols and bioactive peptides on CVD risk, using validated biomarkers and established risk factors for disease. Every step of these dietary intervention studies (from concept to dissemination) has been subject to critical evaluation and governance, in line with the principles of Good Clinical Practice (GCP) and the EFSA evaluation process for the substantiation of a health claim, which in turn has provided the basis to establish this guidance document for the design of future food-based dietary interventions.

Although the primary audience for the toolkit is SMEs considering using or applying for health claims, we realise that this guidance is likely to have value far

beyond the needs of businesses, in particular being a useful resource for students, early-career scientists and others new to research of this nature. The aim of this article is to share the principles of the guide with a broader audience.

Key principles for designing a dietary intervention study

Dietary intervention studies must be well designed in order to yield robust results appropriate for the substantiation of a health claim. Prior to commencing any clinical study, investigators must determine the appropriate study design to answer their scientific question or hypothesis. Aspects such as ethics, the selected population group, the study treatment, the ***outcome*** of interest, as well as the resources available must also be considered.

A well designed study will clearly identify an exposure (*e.g.* amount consumed per day) and the outcome in an objective and quantifiable manner (*e.g.* a reduction in blood pressure) to answer a defined ***study hypothesis***. Multiple stages are involved in the planning of a dietary intervention study (see Fig. 1).

1. Define the study objective and hypothesis

The preparation for a human dietary intervention study begins with the development of a ***research question*** or hypothesis. A study hypothesis is defined as a concept that can be tested in a study or developed as a result of a study. Hypothesis development starts with the collection of background information linked to the area of interest (*e.g.* studies linking polyphenols with cardiovascular health).

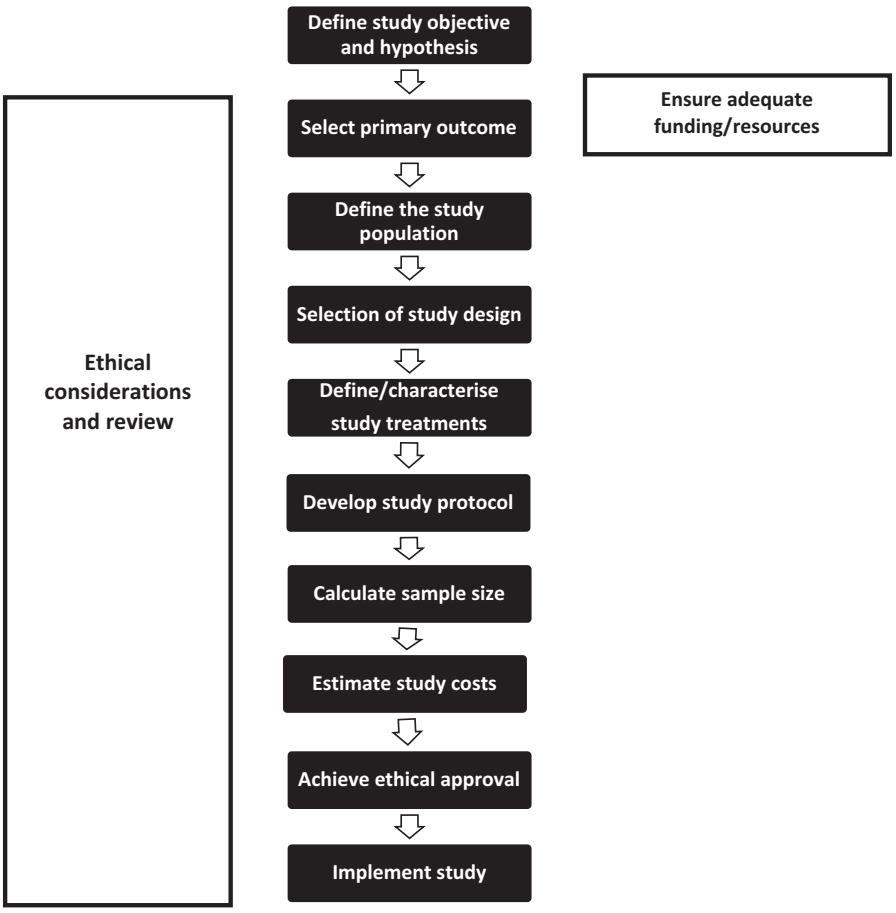

Figure 1 Flow chart of study design for a dietary intervention study.

Example of a study objective from the *BACCHUS* project

The *Flavascular Study* is a dietary intervention to assess the effects of apple-derived polyphenols on CVD risk.

Background from published literature
Apples contain high quantities of flavanols (epicatechin, catechin and procyandin oligomers) and there is good evidence that consumption of these flavanols can reduce CVD risk. However, there are no reported data from RCTs investigating the effects of consuming flavanol-rich apple products on CVD risk. Also, there is an on-going debate as to whether the CVD-protective effects of flavanols are due to the monomers (*e.g.* epicatechin) or the procyanidins, or a combination of both.

Generation of the study objective
To investigate whether consuming apple-derived polyphenols can reduce CVD risk by reducing blood pressure and to differentiate between the cardioprotective effects of epicatechins and procyanidins.

2. Evaluate the scientific evidence

The first step in validating reasoning for the investigation is to complete a thorough review of published literature. This will assist in defining the study rationale, objectives, outcomes and required participant characteristics. Current knowledge in the field may highlight gaps that require addressing or challenges that the proposed study might encounter, which may influence study design and protocol development.

Why are human intervention studies required to establish scientific evidence?

Well designed and conducted dietary interventions/RCTs exist as the gold standard for testing the efficacy and safety of foods and their components. Studies in human volunteers can be classified as either interventional (experimental) or non-interventional (observational; see Fig. 2). The US National Institute of Health defines an interventional study as one in which 'participants receive specific interventions according to the research plan created by the investigators'. In observational studies, there is no intervention; participants are observed and evaluated for exposures and outcomes encountered, as part of the natural course of their lives (Ho *et al.* 2008).

Randomised controlled trials/dietary intervention studies

Dietary intervention studies and randomised controlled trials (RCTs) or clinical trials are terms used interchangeably within this guide. Dietary intervention studies are conducted in exactly the same way as clinical trials testing pharmaceutical agents, but have fewer legislative requirements as dietary intervention studies use food or food-derived components as opposed to *de novo* compounds. Dietary intervention studies are considered the 'gold standard' for testing a hypothesis (*e.g.* evaluating a given treatment) and represent the definitive assessment tool or validation for establishing causal relationships between food components and health and disease risks in humans (Yao *et al.* 2013). The pros and cons of intervention studies are summarised in Table 1.

There are two types of intervention study: clinical trials, where the main aim is to assess the value of new forms of treatment; and field trials, where the objective is to evaluate whether an intervention decreases the risk of disease among disease-free people. For further information on study types, see link from the International Agency on Research for Cancer (IARC) (2015) (www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi-5.pdf).

Observational studies

Observational studies (***cohort*** studies and case–control studies) are useful for understanding relationships between food and health and disease risk. Prospective cohort studies follow a group of individuals into the future to observe whether those that are exposed to certain factors develop a certain disease or specific outcome (*e.g.* the *Framingham Heart study*; www.framinghamheartstudy.org). Observational studies are important for examining associations and for hypothesis generation (AbuMweis *et al.* 2010). Prospective cohort studies are one step below dietary intervention studies in the hierarchy of evidence and while they are of immense value in nutrition and health research, they are not capable of providing definitive evidence of a cause-and-effect relationship. The pros and cons of observational studies are summarised in Table 2.

Animal and in vitro *studies*

While animal studies and *in vitro* models provide valuable underpinning data to elucidate the biological

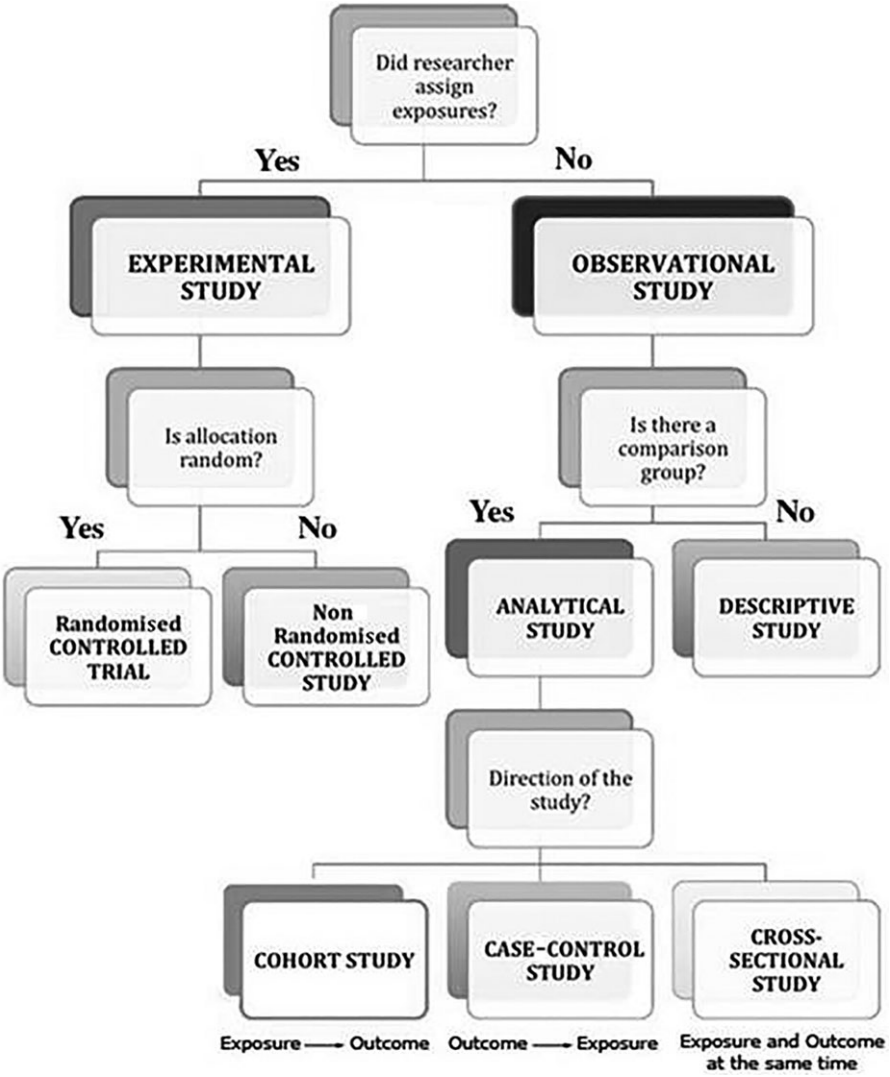

Figure 2 Types of study design [reproduced from Israni (2007) with permission].

plausibility and mechanistic action of a food or food constituent, they are not capable of demonstrating efficacy in humans.

How to identify studies

Systematic approaches for conducting a literature search are outlined in Figure 3 and further guidance can be found at www.bacchus-fp7.eu/Resources/. All studies (including studies reporting null findings) should be considered and used to support the hypothesis or research objective. A review of published studies is also required when seeking ***ethical approval*** to conduct the study.

3. Select the primary study outcome

All study outcomes must be clearly defined, reliable, sensitive, accurate and feasible to measure. They should also should be validated and recognised by the wider scientific community including EFSA.

The primary study outcome should answer the principal research question upon which the design of the study is based. Dietary interventions testing the efficacy of foods should include validated endpoints. For cardiovascular health, examples of study outcomes/surrogate endpoints include blood pressure, endothelial function and blood

Table 1 Main advantages and disadvantages of intervention studies

Advantages	Disadvantages
Investigators have the capacity to evaluate cause-and-effect as well as dose–response relationships	Dietary intervention studies are typically more expensive and time-consuming to perform than observational studies
Generally, dietary intervention studies have stringent selection criteria to ensure subjects are comparable in most respects, thereby reducing confounding influences and isolating effects of the intervention (Besen & Gan 2014)	They require uniquely trained research and medical staff and considerable expertise
By reducing biases and having extensive control over the process, causality of an intervention on a defined outcome can be effectively determined	The applicability of study results to real-world situations may be limited by study population characteristics, procedures implemented or outcomes measured

Table 2 Main advantages and disadvantages of observational studies

Advantages	Disadvantages
Multiple endpoints can be studied at one time	Time-consuming; participants may be followed for many years
A temporal relationship can be examined	Require a large sample size and is expensive to run (resource intensive)
Can be used as evidence to support the rationale for a study	Participants may change their behaviour over time or lose interest in participating in the study

lipids (*e.g.* cholesterol). Outcomes should be clinically significant in terms of health, relate to the health claim being sought and should aim to add independent information about health or a disease risk.

In the *BACCHUS* project, study outcomes for dietary interventions were selected on the basis of their clinical significance for cardiovascular health as well as whether their scientific validity had already been accepted by EFSA.

For further information relating to the selection of cardiovascular outcomes, see *Guidance on the scientific requirements for health claims related to antioxidants, oxidative damage and cardiovascular health* (EFSA 2011) (www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/2474.pdf). Also for study design, see Kendall (2003).

4. Define the study population, and inclusion and exclusion criteria

EFSA requirement for health claim evaluation: inclusion of at least one study in a suitable study group

These are the population group/individuals to whom the results of the study will apply. Justification for the target population should be guided by previously published research (*e.g.* why study populations with a mildly elevated blood pressure were selected instead of adults with a normal blood pressure).

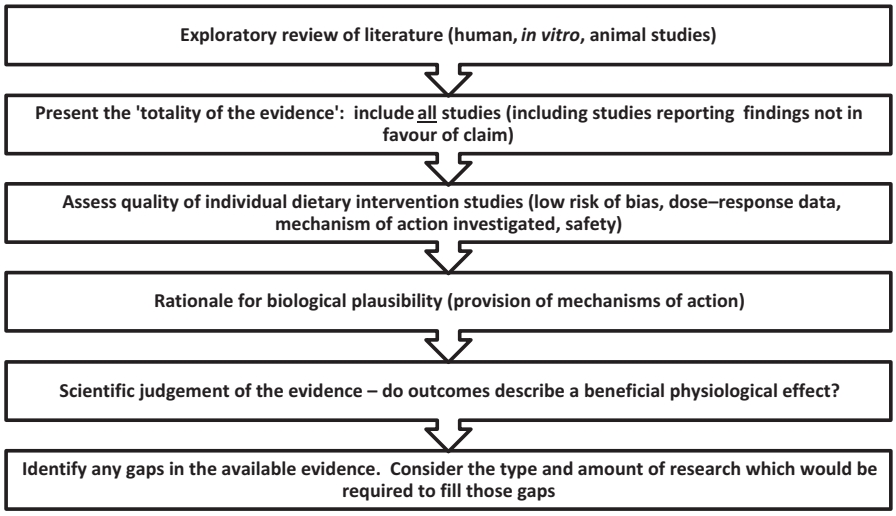

Figure 3 A systematic approach to evaluating scientific evidence (adapted from an EFSA webinar 'Webinar on scientific aspects to consider when preparing health claim application', 10 March 2016). For additional guidance on the evaluation of scientific studies follow the links to EFSA Webinar on scientific aspects to consider when preparing health claim application, 10 March 2016 (www.efsa.europa.eu/sites/default/files/160310-p.pdf) and EFSA General scientific guidance for stakeholders on health claim applications (EFSA 2016) (www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/4367.pdf).

Previous studies will help to define inclusion and exclusion criteria for the recruitment of participants to the study. Inclusion criteria should ensure that the population of interest is clearly defined, while exclusion criteria should not be unduly restrictive as to increase the difficulty in recruiting the required number of study participants by excluding otherwise willing volunteers (AbuMweis *et al.* 2010).

Various factors must be considered when establishing inclusion and exclusion criteria including: ranges for age, gender, ethnicity, bodyweight/BMI, smoking status, medication use, medical history, use of nutritional supplements, alcohol/caffeine consumption, level of physical activity and baseline levels of biochemical markers (*e.g.* blood lipids). The identification of relevant confounding influences is crucial for the development of exclusion criteria, which would deem certain individuals ineligible to participate in the study (AbuMweis *et al.* 2010).

An example of study criteria from the *BACCHUS* project

The *Cardio-Protein Study* is a dietary intervention investigating the effect of egg-white derived peptides on blood pressure (BP) (primary outcome) and CVD risk.

Context: Animal data indicate a BP lowering effect of egg-white derived peptides in rats with elevated BP, but no reduction in BP was observed in rats with normal BP.

Study design: A two-arm, crossover, randomised, double-blind, placebo-controlled RCT.

Study population:
Inclusion criteria

- Caucasian men and women, aged 50–65 years, in good general health (an age group deemed at risk of developing hypertension)
- A high-normal Systolic BP: 130–149 mmHg (volunteers presenting with >150 mmHg were referred to GP as may be hypertensive)
- BMI: 25.0–35.0 kg/m^2
- Informed consent to participate in study provided

Exclusion criteria

- Current smokers
- Diagnosed hypertension, history of CVD, diabetes mellitus (types 1 and 2)
- Depressed or elevated BP (systolic/diastolic: <95/55 mmHg or >150/90 mmHg)
- Medication use: anti-hypertensives, vasodilators, and lipid lowering therapies
- Egg allergy

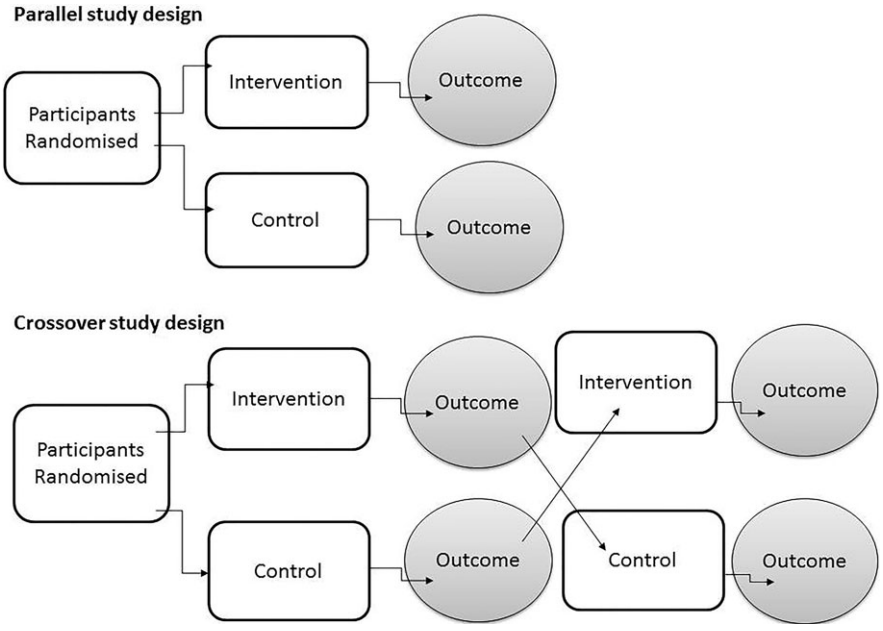

Figure 4 Parallel study design vs. crossover study design

5. Select study design type and randomisation

Dietary interventions studies are generally either a parallel arm or crossover design (see Fig. 4). ***Parallel study design*** involves participants taking part in only one intervention condition/treatment, with different groups of volunteers receiving different interventions (Margetts & Rouse 1997). An example of a parallel design study from the clinical trials registry (Clinicaltrials.gov) can be found here: https://prsinfo.clinicaltrials.gov/trainTrainer/Parallel-Design-Fiction-Manuscript.pdf (Table 3).

A ***crossover study design*** typically involves each participant serving as his/her own control, thus participating in all (two or more) interventions, ideally separated by washout periods. A washout period is a specified amount of time where the study participant does not consume any treatment product. Washout periods are important to minimise potential carry-over effects between study interventions (AbuMweis *et al.* 2010). A carry-over effect is where the response of a treatment from a previous intervention period affects the response of the subsequent intervention period (Margetts & Rouse 1997). The main advantage of a crossover design as compared to a parallel study design is the stronger ***statistical power*** achieved by conducting all treatments on the same individual; this eliminates interindividual variation and thus reduces the overall sample size required (Elbourne *et al.* 2002). An example of a crossover study design from the clinical trials registry (Clinicaltrials.gov) can be found here: https://prsinfo.clinicaltrials.gov/trainTrainer/Crossover-Design-Fiction-Manuscript.pdf (Table 4).

Table 3 Comparison of strengths and limitations of parallel studies

Strengths of a parallel study	Limitations of a parallel study
Studies can be of a shorter duration and results can be generated more swiftly compared with a crossover study	A parallel study can require four to ten times as many participants as a corresponding crossover study in order to achieve the same statistical power (Garcia *et al.* 2004)
No risk of a 'carry-over' effect	The costs associated with implementing a parallel study are significantly greater due to the larger sample size required
Drop-out rates may be reduced due to the shorter study time frame	

Table 4 Comparison of strengths and limitations of a crossover study design

Strengths of a crossover study	Limitations of a crossover study
A smaller sample size is required compared to a crossover study while maintaining statistical power. A smaller sample size reduces recruitment efforts and resources required. This is important for studies that have difficulty recruiting participants due to a narrow selection criteria (AbuMweis *et al.* 2010)	More time-consuming than a parallel study, as each participant must complete all intervention periods
	Participant drop-out rates may be increased due to the prolonged nature of the study
	The need for washout periods, which are necessary to rule out any carry-over effects

Randomisation

Participants are randomised to receive the treatment ('intervention group') or not receive the treatment ('control group'). Both study arms (sometimes there are more than two) are followed in an identical manner and analysed for differences in outcomes.

Random allocation ensures that assignment to a treatment is determined by chance and is not influenced subjectively by investigators or participants. Randomisation is a formal process and simply assigning individuals as recruited to one or other groups is not random. Ideally, randomisation of the study participants should be carried out by an independent person (not involved in conducting the study) to ensure researchers are blinded to the randomisation process and to prevent ***bias***.

Randomisation ensures that any ***confounding variables*** are assigned without bias between groups. If conducted properly, intervention and control groups would be similar in all respects apart from the exposure under investigation [for a review of this area, see Suresh (2011)].

Types of randomisation

- *Simple randomisation* involves randomising a study participant to a treatment based on a single sequence (*e.g.* flipping a coin, heads – *treatment*, tails – *control*). A disadvantage is that unequal samples might be created, especially in studies with a small sample size.
- *Block randomisation* ensures there are equal numbers in both the treatment and control groups through the creation of blocks of samples (*e.g.* n = 4, 6 or 8) by the researcher. This method is used to ensure a balance in sample size across groups over time. Blocks are small and balanced with predetermined group assignments, which keeps the numbers of participants in each group similar at all times.

• *Stratified randomisation* involves the creation of strata, which groups individuals according to a certain characteristic (*e.g.* age). This stratified randomisation method controls for the possible influence of covariates that could jeopardise the conclusions of the research. For example, age may be a confounding variable and influence the outcome of the research. Stratified randomisation can balance the control and treatment groups for age or other identified covariates.
• *Covariate adaptive randomisation* is when a new participant is sequentially assigned to a particular treatment group by taking into account specific covariates and the previous assignments of other participants.

6. Select and characterise study treatments

When selecting suitable study treatments, subject-related factors such as age, gender, health status, as well as intervention characteristics, such as the form and vehicle of the food constituent under investigation, must be considered (AbuMweis *et al.* 2010). In addition, the dose, frequency and diurnal timing of intake of the food/food component within the dietary intervention also merit review.

Selection of the study treatment product

(1) The food/constituent investigated should comply with the specifications of the food/constituent for which the claim is proposed.

(2) If test and control foods differ in characteristics other than the food/constituent, it could affect the claimed effect (*e.g.* dark vs. white chocolate).

EFSA Journal 2016; 14(1):4367 p13-14

Inclusion of a comprehensive food ***characterisation*** of study products is essential for the substantiation of a health claim; this includes data on the physical and chemical attributes, nutritional composition, manufacturing processes, batch variability and safety data (see Fig. 5). Without adequate characterisation of the active constituent, a claim application is likely to be rejected (*i.e.* receive a negative opinion) regardless of the quality of the scientific evidence. Additional practical aspects, which should be considered, include shelf-life stability, microbiological safety, nature of the product, consumer acceptability, suitability as a study treatment, packaging, logistics and costs. The importance of carefully considering study products is discussed in more detail in the *Best Practice Guide for Health Claims* (www.bacchus-fp7.eu/) and is stressed in the recently updated EFSA guidance for health claims applicants (EFSA 2016).

In addition, information on the digestive and absorptive processes and the bioavailability of the food component within the chosen food matrix should also be considered. Bioavailability is the degree to which a food/food constituent becomes available to the target tissue (*i.e.* intestines for absorption) after administration. Assessing the bioavailability of the active food constituent within the given food matrix, perhaps using animal studies, is an important exercise prior to

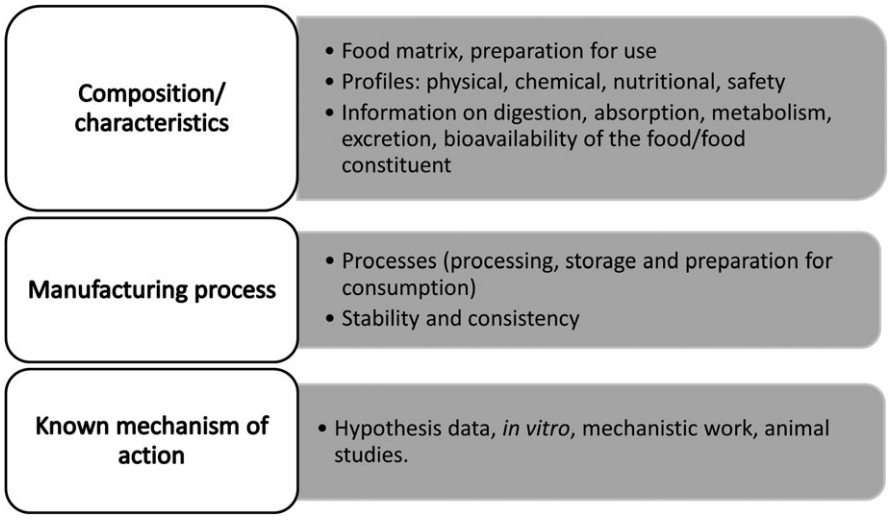

Figure 5 Characterisation of study treatments.

utilising it in a human study. Finally, further *in vitro*/ mechanistic work or animal studies may be required to demonstrate potential biological mechanisms of action.

Selection of a dosage regimen

Studies are designed to include either a single dose or multiple (dose–response) treatment arms. Study doses should be clearly defined, relevant to the study outcome and the health claim and based on previous scientific evidence where possible. Choosing the amount of the active food ingredient and the frequency of consumption should involve consideration of efficacy, safety and the feasibility of incorporating a minimum effective amount into usual consumption patterns.

The maximum effective amount of an active food ingredient, above which there is no further improvement in efficacy should be based on a previously established dose–response relationship. Dose–response ranges should be wide enough to demonstrate the biological effect (AbuMweis *et al.* 2010). Most importantly, the dosing regimen should answer the following question – 'can such a dose/amount of food constituent be reasonably consumed within the context of a balanced diet?' – this is an important consideration when compiling evidence to support a health claim.

Prior to conducting a human study, it is critical to establish a comprehensive safety profile for the food/ food component under review. Dose–response extrapolation from animal or observational study data should be carried out with caution (AbuMweis *et al.* 2010). Safety trials carried out in animal or *in vitro* studies include tests for absorption, metabolism, excretion and the kinetics of the compound. Once safety has been established, a dietary intervention study can be used to further corroborate the absence of any ***adverse effects***.

The blinding of study treatments

Use of a ***placebo*** control: the placebo product should be similar in appearance and taste to the treatment but void of the active ingredient. The use of a placebo within a study enables blinding of both participants and researchers. Blinding reduces the likelihood that the behaviours of subjects or investigators could influence the results of the study. In a single-blinded study, subjects are unaware of their treatment status; whereas in a ***double-blinded*** study, both the investigators and the subjects are unaware of which intervention receive. The creation of a true placebo in food-based intervention studies can be particularly challenging, as it can be difficult to mask many food constituents due to their unique taste, sensory characteristics or colour. A further obstacle unique to dietary studies is that the addition of a food constituent may cause an imbalance in dietary intakes of macro-/micronutrients between groups. Thus, efforts should be made to match the control diet with the experimental diets in terms of total nutrient and energy intake.

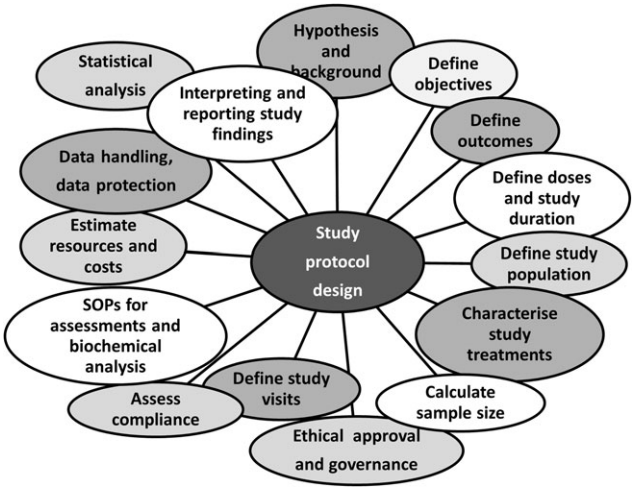

Figure 6 Components of a study protocol (SOPs: standard operating procedures)

7. Develop the study protocol

A 'study protocol' document describes in detail the plan for conducting the study. The study protocol explains the purpose and function of the study, as well as how to implement the study (see Fig. 6).

A study protocol includes a detailed description of the hypothesis, objective, outcomes, the study products, the study population, inclusion and exclusion criteria, calculation of sample size with scientific justification, plans for ethical governance, a detailed description of study visits and estimation of resources and equipment required. In addition, during this stage, standard operation procedures (SOPs) and questionnaires are prepared for assessments at study visits, data handling and protection, sample collection and bio-banking, laboratory analysis and statistical analysis (Sakpal 2010). The use of a SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) check-list ensures that all key study design elements have been considered. SPIRIT is an international initiative that aims to improve the quality of clinical trial protocols by defining an evidence-based set of items to address in a protocol. For further information see www.spirit-statement.org/. The phases within a dietary intervention study are outlined in Figure 7.

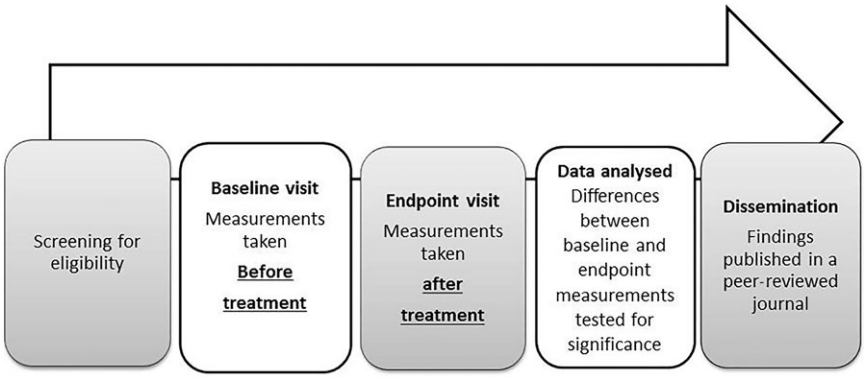

Figure 7 Example timeline of a dietary intervention study (parallel design)

Estimation of study costs

The development of a well-defined protocol aids the estimation of resources, training and equipment required. An adequate funding source should be secured in advance of study planning, with potential to accommodate unanticipated cost overruns. It is critical that human and infrastructural resources are planned well in advance of the study being launched.

Selection of study duration

Appropriate study durations should confirm that efficacy is demonstrated and, if desired, a sustainability of effect is established. Prolonged interventions over several months can be linked with higher costs, greater dropout rates, diminished compliance and logistical hurdles – all factors which render extended interventions prohibitive. Aside from budget and resource availability, other key factors for the selection of optimal study length include the amount of time required to reach a stable effect on endpoint/outcome measures (*i.e.* to capture as significant a biological effect as is feasible) and the acceptability of the dietary treatments and dietary restrictions over a prolonged period (AbuMweis *et al.* 2010).

Assessment of participant compliance with a study protocol

Participant compliance with study treatments (*e.g.* consumption of the test product) is essential to ensure the validity of results. Monitoring participant adherence to dietary treatments is important to ensure that a lack of compliance is not the cause of any observed effect (Kehoe *et al.* 2009). Dietary assessments, compliance questionnaires, pill-counting, and if possible, measurement of specific blood, urine and faecal biomarkers of exposure can be used to assess compliance.

Control of background diet

For studies involving dietary intervention, intakes of certain foods may be restricted for the duration of the study. Several regimens may be used to control for background diet ranging from free-living (no controls imposed) to a partially controlled or a metabolically controlled regimen, where participants are required to consume their dietary interventions under the daily supervision of the research team which can be restrictive and labour intensive for both the study participant and researcher (Margetts & Rouse 1997).

8. Calculate sample size

Determining the appropriate sample size (*i.e.* the minimum number of study participants required to detect a scientific effect when one exists) is a vital element of study design and is based on the primary study outcome. Incorrect sample size estimations may lead to serious flaws in study design and ultimately failure to correctly identify efficacious treatment. The components required to calculate sample size for a randomised dietary intervention are the ***statistical power*** and the level of ***statistical significance***, the likely size of the treatment effect sought and its variability (Noordzij *et al.* 2010). This information can be established using data from previous studies or by review of relevant published studies in the area. Crossover and parallel designs use a different formula for sample size estimation; thus, it is desirable to consult an experienced statistician when estimating sample size (Sakpal 2010).

Notes on estimating sample size

In an intervention, the default assumption is 'no effect'. Thus, the intervention seeks to reject the default or *Null Hypothesis* (H_0), which states that

there is no relationship between the treatment and observed effects, and accept the *Alternate Hypothesis* (H_1), which states there is a relationship between treatment and observed effects.

A *Type I error* occurs when the null hypothesis is wrongly rejected, stating that there is a relationship between the treatment and observed effects when this is untrue ('false positive').

A *Type II error* occurs if the null hypothesis is incorrectly retained when a relationship between the treatment and observed effects exists ('false negative'). Errors in sample size calculation may involve *sample size under-estimation*, where the sample size selected is less than that required to detect an effect of treatment on the specified outcome when it exists (*i.e.* insufficient statistical power). In other words, under-estimation of sample size might result in a type II error (failure to see a statistically significant effect where one would exist had there been sufficient statistical power).

The level of statistical significance is the probability of observing an effect given that the null hypothesis is true and is typically set at 5% (0.05) or less. The risk of a Type I error is inversely proportional to the level of statistical significance. Statistical power is determined by sample size and should not be less than 80%. The risk of a Type II error is directly proportional to sample size.

Information about expected response differences between the treatment and control groups can be obtained from previous studies conducted using a similar treatment. If there have been no previous studies, a feasibility study may be conducted. A feasibility study is a study conducted on a smaller scale to guide decisions on appropriate sample size and study design for the main study (Arain *et al.* 2010). Other factors that should be considered include the *clinical importance/ meaningful difference*, which require defining the difference between test and reference treatments that could be considered clinically meaningful. This threshold figure is, at times, not readily available and should be decided based on clinical judgment (Sakpal 2010).

Example power calculation

Primary outcome measure for this parallel study is LDL cholesterol (LDL-C).

Background: Data from previous studies at the Institute of Food Research (*BASH* study 12/EE/0313 and The *DVH* study 09/H0311/96) suggest that baseline values of LDL-C for a population not dissimilar to that which we plan to include ranging from 2.05 to 6.16 mmol/l (n = 57). Data from the literature (two studies) investigating the effects of anthocyanins on LDL-C, report mean levels ranging from 3.10 to 3.36 mmol/l (n = 15) and from 3.13 to 3.44 mmol/l (n = 25).

(1) The sample size calculation will assume a baseline LDL-C of 3.23 mmol/l.

The standard deviation of longitudinal differences in LDL-C reported in these studies ranged between 0.41 and 0.94 mmol/l. The lower value is consistent with the pooled standard deviation in the longitudinal difference in LDL-C of 0.41 mmol/l reported in another study on flavan-3-ols.

(2) The sample size calculation will assume a standard deviation of the paired difference in LDL-C of 0.41 mmol/l

The final component of a power calculation refers to the effect size that one would predict. Previous studies displayed reductions of 8.6%–21.2%, whereas other studies showed reductions in LDL-C associated with anthocyanins and flavan-3-ols of 13.8% and 5.2% respectively. Doses of test products vary enormously and so a conservative estimate of a 5% decrease (=0.16 mmol/l) will be used, considering our dose of 50 mg of anthocyanin/day. This is less than 10% of the typical reduction in LDL-C associated with statins (1.8 mmol/l) reported in a meta-analysis.

(3) The sample size calculation will assume a paired difference in LDL-C of 0.162 mmol/l.

Sample size has been powered on the primary outcome (LDL-C) using a one-sided paired comparison for 5% significance and 80% power – the number of participants required is 42.

Sample size calculator: A free tool for calculating the sample size for human studies developed by Massachusetts General Hospital, Biostatistics Centre, Boston, US, can be found at www.hedwig.mgh.harvard.edu/sample_size/size.html.

Dropouts: While sample size estimation will provide the number of participants required to complete a study to achieve the desired statistical significance for a given hypothesis, in practice, more participants (10–20%) should be enrolled to account for potential dropouts. A participant may decide to discontinue with a study for a variety of reasons, such as adverse response to the treatment/placebo, illness or personal reasons (*e.g.* time commitment). If a participant decides to discontinue with a study, they are considered a 'study dropout'. Study dropouts can reduce the statistical power of the study and, therefore, must be accounted for when designing a study. One way to overcome this is through the recruitment of additional participants. Sample size calculations can be adjusted to account for dropouts (Sakpal 2010).

Adjusted sample size calculator

n = sample size required; d = dropout rate; N1 = adjusted sample size

$N1 = n/(1 - d)$

Note – the prediction of a potential dropout rate should be realistic as the over-recruitment of participants to participate in a study protocol may violate ethical considerations.

9. Consider ethical issues and governance

All dietary intervention studies in humans must achieve ethical approval prior to the screening and recruitment of participants. The ethical review of a research study is an important way of protecting individuals from harmful or poorly designed studies and ensuring that participants are properly informed of the nature of the research to which they are consenting. A detailed study protocol providing comprehensive data about the proposed study is submitted to a relevant institutional ethical governing body for a thorough review and, on these grounds, a decision is made as to whether or not a study can proceed. All study protocols should be designed and implemented in accordance with the principles of GCP and incorporate a monitoring system for the assessment of quality control throughout the study. An introductory tutorial on ethics in research is at www.cirt.gcu.edu/research/developmentresources/tutorials/ethics.

Information usually required for an ethical submission for a dietary intervention study:

- governance details, including CV of the Principal Investigator;
- details of indemnity, sponsorship and regulatory and institutional approval;
- background/literature review and rationale for the study hypothesis;
- study objectives and outcomes;
- detailed study design;
- description of the study population, inclusion and exclusion criteria;
- recruitment strategies and materials to be used for recruitment;
- description of study treatments (including safety data);
- description of participant contacts, screening and study visits;
- details of roles and responsibilities of study participants;
- clarification that study will be conducted in accordance with the principles of GCP, description of GCP compliant systems, processes and training;
- plan for biobanking and biochemical analysis;
- data protection plan and and statistical analysis;
- roles and responsibilities of the study team;
- safety issues/foreseen risks/safety monitoring (*e.g.* biochemical) and adverse events;
- study monitoring and data monitoring;
- all questionnaires, materials and participant information leaflets, consent forms and permitted contact scripts, either by phone, SMS or email.

Good clinical practice

GCP by the International Conference on Harmonization (ICH) is an international standard for the design, conduct, performance, monitoring, auditing, recording, analyses and reporting of human clinical trials (including dietary intervention studies). Researchers must protect the health, privacy and dignity of study participants through the use of an appropriate study design that is clearly defined and demonstrates that all ethical considerations have been made. All study participants must provide informed consent and participation must be voluntary. GCP provides assurance that the data and reported results are credible and accurate and that the rights, integrity and confidentiality of study participants are protected.

Informed consent is a process by which a volunteer voluntarily confirms his or her willingness to

participate in a trial after having been informed of all aspects of the trial. Informed consent is documented by means of a signed and dated consent form.

When designing consent forms, considerations include the following:

- how measurements for the proposed outcome(s) will be achieved; for biological samples, how these will be stored, used and destroyed subsequently (*e.g.* blood sampling);
- if there are any previously reported adverse effects;
- and evidence (*e.g.* a signature) that a participant fully comprehends the information provided.

Detailed participant information sheets must accompany consent forms. Examples and templates for informed consent from the World Health Organization (WHO) can be found at www.who.int/rpc/research_ethics/informed_consent/en.

A Summary of the Principles of Good Clinical Practice (GCP) by the International Conference on Harmonization (ICH)

(1) Clinical trials should be conducted in accordance with the ethical principles that have their origin in the Declaration of Helsinki*, and that are consistent with GCP and the applicable regulatory requirement(s).
(2) Before a trial is initiated, foreseeable risks and inconveniences should be weighed against the anticipated benefit for the individual trial subject and society. A trial should be initiated and continued only if the anticipated benefits justify the risks.
(3) The rights, safety, and well-being of the trial subjects are the most important considerations and should prevail over interests of science and society.
(4) The available nonclinical and clinical information on an investigational product should be adequate to support the proposed clinical trial.
(5) Clinical trials should be scientifically sound, and described in a clear, detailed protocol.
(6) A trial should be conducted in compliance with the protocol that has received prior institutional review board (IRB)/independent ethics committee (IEC) approval/favourable opinion.
(7) The medical care given to, and medical decisions made on behalf of, subjects should always be the responsibility of a qualified physician or, when appropriate, of a qualified dentist.
(8) Each individual involved in conducting a trial should be qualified by education, training, and experience to perform his or her respective task(s).
(9) Freely given informed consent should be obtained from every subject prior to clinical trial participation.
(10) All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation and verification.
(11) The confidentiality of records that could identify subjects should be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s).
(12) Investigational products should be manufactured, handled, and stored in accordance with applicable good manufacturing practice (GMP). They should be used in accordance with the approved protocol.
(13) Systems with procedures that assure the quality of every aspect of the trial should be implemented.

For further information see *International Conference on Harmonization, Harmonised Tripartite Guidelines: Guidelines for Good Clinical Practice 1996 / Under revision 2015* (www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf and more recent revisions available at the European Medicines Agency: www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/08/WC500191488.pdf).

* The Declaration of Helsinki (World Medical Association (WMA) 2013).

Reporting of adverse events

An adverse event is an unexpected and unfavourable side effect of the intervention and may be either treated as a non-serious or serious adverse event. All adverse events and serious adverse events must be documented and reported to the institutional ethics review body and decisive actions taken.

A serious adverse event is one in which the study treatment results in death, a life-threatening event, requires inpatient hospitalisation or prolongation of existing hospitalisation, or results in persistent or significant disability/incapacity or results in a congenital anomaly/birth defect (in pregnancy studies) (ICH 1996).

All suspected adverse events should be recorded and documented. In the case of a suspected serious adverse event, the investigator must notify the sponsor, which can be an individual, an institution, company or organisation that takes responsibility for the management and financing of the research but is not involved in conducting the research. The investigator must notify the sponsor within 24 hours if a serious adverse event occurs. The sponsor should then report this without delay to the external institutional ethics review board. Further information about the reporting of adverse events can be found at www.ec.europa.eu/health/files/eudralex/vol-10/2011_c172_01/2011_c172_01_en.pdf.

Additional considerations

- Every clinical trial must be registered in a publicly accessible database prior to recruitment of the first study participant (*e.g.* www.clinicaltrials.gov/).
- A data protection protocol should be in place prior to commencing participant recruitment.
- Ethical considerations are a priority and warrant continuous review throughout the duration of the study. It is the responsibility of the research investigators to report to the relevant ethical authority the occurrence of any adverse events, as well as any amendments to the study protocol.

Good manufacturing practice

If a food product/constituent is developed with the intention for use in a dietary intervention study, then international standards for good manufacturing practice (GMP) would apply. GMP ensures quality assurance by confirming that all processes in the manufacturing of the product are accurately validated and documented.

WHO guidelines for GMP can be found at www.who.int/biologicals/vaccines/good_manufacturing_practice/en/ and ICH-GMP guidelines at www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q7/Step4/Q7_Guideline.pdf.

10. Analyse data and report study findings

The sequence of events for the generation of study findings is presented in Figure 8. The *Consolidated Standards of Reporting Trials* or *CONSORT* statement is an evidence-based minimum set of recommendations which aim to improve the reporting of study findings for RCTs and dietary interventions. Many journals specify the use of a *CONSORT* statement as a condition for publication. See here for further information: www.consort-statement.org/.

All scientific outputs must be transparent, understandable and reproducible with regard to the data, methods of analysis and assumptions that are used in the risk assessment process. Any statistical difference reported should be interpreted in the light of its biological relevance. For health claim substantiation, demonstration of statistical significance (at $P < 0.05$) of the food's effect is pivotal.

Expression of uncertainty and variability in risk estimates should be quantified to the extent that is scientifically achievable. There may be differences in risk due to variability among individuals, populations, species or ecosystems (*i.e.* confounding influences). It is important to identify the most influential contributors to variability in risk and to control for them in the statistical analyses. Finally, where possible, harmonised assessment terminology should be used, preferably based on internationally accepted terminology.

EFSA (2014) has published guidance on the interpretation and reporting of study findings (www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/3908.pdf).

Figure 8 Linear sequence for data analysis and reporting of study findings

Conclusion

Well-designed and conducted dietary interventions and trials are the gold standard for testing the efficacy and safety of functional foods and their constituents (AbuMweis *et al.* 2010). These studies are also at the top of the hierarchy for informing decisions related to health claim substantiation, as well as informing regulatory authorities who develop nutrient recommendations and dietary guidelines.

A dietary intervention study should:

- use comprehensively characterised study treatments;
- be short enough to optimise subject compliance and be cost-effective, while lengthy enough to ensure biological efficacy and to avoid chance findings;
- select validated, appropriate outcome variables for the claimed effect;
- be conducted in a suitable study group, sufficiently powered and with all care and attention to avoid high participant dropout rates;
- adhere to GCP guidelines and fully consider regulatory and ethical legislation;
- and be transparent in the analysis, interpretation and reporting of study findings.

In addition, the development of a well-defined study protocol requires careful planning and adequate resources, with appropriately qualified and indemnified staff. An adequate budget should be established from the outset to cover all costs and accommodate any potential overspends. Paying attention to all of these elements is crucial to the design of quality dietary studies that can reliably evaluate relationships between food and health.

This guide has been developed with the needs in mind of those collating evidence to support a health claim application. However, it is also a useful summary of the necessary steps and principles for students, early-career scientists and others new to research of this nature.

Acknowledgements

This work is supported by the *BACCHUS* project which has received funding from the EU Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No 321090.

Glossary of useful terms

Adverse effect: A harmful effect different to the expected treatment effect.

Bias: The encouragement of one outcome over another. The presence of bias reduces the validity of a study.

Cause and effect: It (also referred to as causation or causality) is the action or efficacy that connects one process (the *cause*) with another (the *effect*), where the first is understood to be partly responsible for the second and the second is dependent on the first.

Characterisation: A comprehensive description of the various components of an item.

Confounding variable: An extraneous variable in a study design that correlates with both the dependent and independent variables; this variable which has not been accounted or controlled for may affect either the treatment or the outcome, possibly leading to false results. Examples of potential confounding variables include age, gender, body mass index and socio-demographic characteristics.

Cohort: A group of individuals with certain characteristics that are studied over a period of time to examine the incidence of a certain disease, all cause deaths or another outcome.

Double-blinded study: Both the participants and the researchers are unaware of the treatment that the participant has been randomly allocated to.

Efficacy study: Refers to an intervention study (in humans; in animals) which investigates the relationship between the food/constituent and the claimed health effect.

Ethical approval: Before the conduction of a study, an application, outlining the study protocol, is submitted to the relevant institutional governing board for ethical review and approval to proceed. A study must be conducted in accordance with the principles of GCP and guidelines established by the Declaration of Helsinki. This independent ethics board will evaluate and decide whether study participants will be adequately protected, in terms of their rights and safety if they choose to enrol in the proposed study.

Food/food constituent: Refers to a food category, a food or a food constituent (*e.g.* a nutrient or other substance) or a fixed combination of nutrients/other substances.

Outcome: An outcome in research is a result or finding from a study. A primary outcome is the main measurement (*e.g.* blood pressure) that is being investigated and upon which the study deign and sample size estimation is based. A secondary outcome is a measurement in which the study may have not been originally designed for but may have observed an effect (*e.g.* blood cholesterol).

Pertinent study: It is a robust and valid study from which scientific conclusions can be drawn.

Placebo: A study product that is similar to the treatment but void of the active ingredient.

Statistical power: The power or statistical power of a study is the likelihood of finding an effect of the intervention if present. A well-powered study is designed to observe any effects that may occur as a result of the intervention and not by chance (Sink & Mvududu 2010).

Statistical significance: To determine whether an observed effect is caused by a variable (*e.g.* active treatment) and not by chance. The *P*-value or calculated probability is the probability of finding the observed results when the null hypothesis (H_0) of a study question is true. The null hypothesis is usually a hypothesis of 'no difference' (*e.g.* no difference between blood pressures in group A and group B). Define a null hypothesis for each study question clearly when designing a study. A statistical test usually considered significant if the *P*-value is <0.05 indicating a relationship between the treatment and the observed effect. Most researchers refer to statistically significant as $P < 0.05$ and statistically highly significant as $P < 0.001$ (less than one in a thousand chance of being incorrect).

Study hypothesis: A concept that can be tested in a study or developed as a result of a study.

References

AbuMweis SS, Jew S & Jones PJH (2010) Optimizing clinical trial design for assessing the efficacy of functional foods. *Nutrition Reviews* **68**: 485–99.

Arain M, Campbell MJ, Cooper CL *et al.* (2010) What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Medical Research Methodology* **10**: 1–7.

Besen J & Gan SD (2014) A critical evaluation of clinical research study designs. *Journal of Investigative Dermatology* **134**: e18.

Buttriss JL (2015) Nutrition and health claims in practice. *Nutrition Bulletin* **40**: 211–22.

Elbourne DR, Altman DG, Higgins JPT *et al.* (2002) Meta-analyses involving cross-over trials: methodological issues. *International Journal of Epidemiology* **31**: 140–9.

European Commission (EC) (2007) EC Regulation No 1924/2006 on nutrition and health claims made on foods. *The Official Journal of the European Union* **L12/18**: 1–16.

European Food Safety Authority (EFSA) (2011) Guidance on the scientific requirements for health claims related to antioxidants, oxidative damage and cardiovascular health. *EFSA Journal* **9**: 2474. doi:10.2903/j.efsa.2011.2474.

EFSA (European Food Safety Authority) (2014) Guidance on Statistical Reporting. *EFSA Journal* **12**: 3908. doi:10.2903/j.efsa.2014.3908.

European Medicines Agency: ICH E6 (R1) Good clinical practice, Revision 2 (2015). http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/08/WC500191488.pdf (accessed 30 September 2016).

EFSA (European Food Safety Authority) (2016) General scientific guidance for stakeholders on health claim applications. *EFSA Journal* **14**: 4367. doi:10.2903/j.efsa.2016.4367.

Garcia R, Benet M, Arnau C *et al.* (2004) Efficiency of the cross-over design: an empirical estimation. *Statistics in Medicine* **23**: 3773–80.

Ho PM, Peterson PN & Masoudi FA (2008) Evaluating the evidence: is there a rigid hierarchy? *Circulation* **118**: 1675–84.

IARC (International Agency for Research on Cancer) (2015) France (www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi-5.pdf) (accessed 20 June 2016).

ICH (International Conference on Harmonization) (1996) Harmonised Tripartite Guidelines. Guidelines for Good Clinical Practice. Available at: www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf (accessed 21 June 2016).

ICH (International Conference on Harmonization) (2000) Harmonised Tripartite Guidelines. Guidelines for Good Manufacturing Practice guide for active pharmaceutical ingredients. Available at: www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q7/Step4/Q7_Guideline.pdf (accessed 21 June 2016).

Israni RK (2007) Guide to Biostatistics. Med Page Today. Available at: www.medpagetoday.com/lib/content/Medpage-Guide-to-Biostatistics.pdf (accessed 21 June 2016).

Kehoe SH, Chheda PS, Sahariah SA *et al.* (2009) Reporting of participant compliance in randomized controlled trials of nutrition supplements during pregnancy. *Matern Child Nutr* **5**: 97–103.

Kendall JM (2003) Designing a research project: randomised controlled trials and their principles. *Emergency Medical Journal* **20**: 164–8.

Margetts BM & Rouse IL (1997) Experimental studies: clinical trials, field trials, community trials & intervention studies. In: *Design Concepts in Nutritional Epidemiology*, 2nd edn (BM Margetts & M Nelson eds), pp. 415–39. Oxford University Press: New York.

Noordzij M, Tripei G, Dekker FW *et al.* (2010) Sample size calculations: basic principles and common pitfalls. *Nephrology Dialysis Transplantation* **25**: 1388–93.

Sakpal TV (2010) Sample size estimation in clinical trials. *Perspectives in Clinical Research* **1**: 67–9.

Sink CA & Mvududu NH (2010) Statistical power, sampling, and effect sizes: three keys to research relevancy. *Counselling Outcome Research and Evaluation* **1**: 1–18.

Suresh K (2011) An overview of randomization techniques: an unbiased assessment of outcome in clinical research. *Journal of Human Reproductive Sciences* **4**: 8–11.

WMA (World Medical Association) (2013) Declaration of Helsinki. *Bulletin of the World Health Organization* **79**: 373–4.

Yao CK, Gibson PR & Shepherd SJ (2013) Design of clinical trials evaluating dietary interventions in patients with functional gastrointestinal disorders. *American Journal of Gastroenterology* **108**: 748–58.