

Information on establishing food-health relationships for general level health claims

**(as described in Schedule 6 in the
Australia New Zealand Food Standards Code)**

Food Standards Australia New Zealand

**Version 1.2
May 2016**

Disclaimer

This document provides general information only and may be subject to change at any time without notice.

To the extent permitted by law, Food Standards Australia New Zealand (FSANZ) disclaims and shall not be liable for any injury, loss, damages, costs and expenses arising in any way, including by way of negligence, from or in connection with any information provided or omitted, or from anyone acting or refraining to act in reliance on the information in this document. Reference to FSANZ includes a reference to any contractor, agent or employee of FSANZ.

Any person relying on the information in this document should seek their own independent legal advice in relation to any queries they may have regarding obligations imposed under the *Australia New Zealand Food Standards Code* (the Code) and relevant food legislation and other applicable laws. These laws include: *Food Standards Australia New Zealand Act 1991* (Cth) (the FSANZ Act); the *Competition and Consumer Act 2010*; the *Imported Food Control Act 1992*; and state and territory fair trading Acts and food Acts. In New Zealand, relevant legislation includes the *Food Act 2014* and *Fair Trading Act 1986*. Legal requirements may change as government regulations are made or changed or interpreted by the courts.

FSANZ has taken great care to ensure that the material provided in this document is as correct and accurate as possible. However, FSANZ makes no warranty with respect to the accuracy, reliability or completeness of the material contained in this document or that the document, if used or relied on, will ensure compliance with the relevant requirements of the FSANZ Act and/or the Code.

FSANZ makes material in the document available on the understanding that you will exercise your own skill, care and judgment with respect to its use and you will carefully evaluate the accuracy, currency, completeness and relevance of the material for your purposes.

Links to third party websites

This document contains links to websites that are external to FSANZ. These links are provided for information purposes only. FSANZ has no direct control over the content of the linked sites, or the changes that may occur to the content on those sites. FSANZ is not responsible for the accuracy, legality, content or reliability of the external site or for that of subsequent links, or for any loss or inconvenience arising from the use of such sites and links. Providing links to external websites does not constitute an endorsement or a recommendation of any material on those sites or of any products, services or opinion offered by, from or through those sites. Users of links provided by this document are responsible for being aware of which organisation is hosting the website they visit and for making their own decisions about the relevance or accuracy, currency and reliability of information found on those sites.

Food standards in Australia and New Zealand

The Australian and New Zealand food standards system is governed by legislation in the states, territories, New Zealand, and the Commonwealth of Australia; including the FSANZ Act. The FSANZ Act sets out how food regulatory measures are developed. It establishes FSANZ as the agency responsible for developing and maintaining the Code.

Responsibility for enforcing the Code in Australia rests with authorities in the states and territories; the Commonwealth Department of Agriculture and Water Resources for imported food; and with the Ministry for Primary Industries in New Zealand.

Acknowledgement

FSANZ would like to acknowledge Health Canada for permission to reproduce material from their document titled *Guidance document for preparing a submission for food health claims* (2009).

Copyright

FSANZ supports and encourages the dissemination and exchange of information.

The Licence

Material in this document is provided under a **Creative Commons Attribution 3.0 Australia (CC BY 3.0 AU) Licence**, except for:

- the Food Standards Australia New Zealand logo; and
- third party material.



The details of the licence conditions and the full legal code for the CC BY 3.0 AU Licence are available on the [Creative Commons website](https://creativecommons.org/licenses/by/3.0/au/).

What the Licence allows

In summary, the CC BY 3.0 AU Licence allows you to share the material in this publication in any medium or format; and adapt the material, for commercial and non-commercial purposes, subject to the following restriction:

Attribution: you must attribute the work in the following manner:

© Food Standards Australia New Zealand.

The attribution must not, in any way, suggest that FSANZ endorses you or your use of the work.

Third party material

To the extent that material in this document contains material in which copyright is owned by a third party, the CC BY 3.0 AU Licence would not apply to such third party material and, if you wish to re-use the third party material, you may have to seek permission from the copyright owner.

An electronic version of this document is available on the FSANZ website at www.foodstandards.gov.au and, for New Zealand, www.foodstandards.govt.nz.

For more information email information@foodstandards.gov.au.

Contents

- Disclaimeri
- Acknowledgement ii
- 1 Introduction 1
 - 1.1 Purpose of this document..... 1
 - 1.2 Requirements for notifying FSANZ of an established food-health relationship..... 1
 - 1.3 Provision of records to a relevant authority..... 2
- 2 Overview of substantiating a food-health relationship..... 2
 - 2.1 Substantiation 2
 - 2.2 Preparation of the systematic review report..... 3
 - 2.3 Structure of the systematic review report..... 3
- 3 The systematic review 4
 - 3.1 Reference material for conducting a systematic review 4
 - 3.2 Description of the food-health relationship..... 6
 - 3.3 Retrieval of scientific evidence – systematic review based on the original literature only..... 8
 - 3.4 Tabulation of data from the final list of included studies.....12
 - 3.5 Assessment of study quality13
 - 3.6 Assessment of the body of evidence and conclusion.....14
 - 3.7 Applicability to Australia and New Zealand.....16
 - 3.8 Updating an existing systematic review18
- References.....20
- Glossary22
- Appendix 1: Template for a PRISMA flow diagram documenting filtering of studies retrieved in a literature search25
- Appendix 2: Examples of table layout for showing study characteristics26
- Appendix 3: Examples of quality appraisal tools28
- Appendix 4: The GRADE system.....30
- Appendix 5: Revision History32

1 Introduction

1.1 Purpose of this document

This document provides a general overview for food businesses wishing to establish a relationship between a food or property of food and a health effect (food-health relationship) by a process of systematic review for the purpose of making a general level health claim¹.

All the requirements for making a general level health claim on a food label or in an advertisement are set out in Standard 1.2.7 – Nutrition, Health and Related Claims in the *Australia New Zealand Food Standards Code* (Code).

Food businesses wishing to make a general level health claim can base their claim on a food-health relationship that is either:

- pre-approved by FSANZ as listed in the table to section S4—5 in the Code, or
- established in accordance with requirements set out in Schedule 6.

This document outlines scientific best practice for undertaking a systematic review as described in Schedule 6. The text presented in shaded boxes presents the requirements from Schedule 6 – Required elements of a systematic review.

For the purposes of this document, *substantiation* refers to the process of evaluating the evidence for a food-health relationship to underpin a general level health claim as required under paragraph 1.2.7—18(3)(b).

For the purposes of this document, an *established food-health relationship* refers to a food-health relationship for which evidence has been examined using the substantiation process and a reasonable conclusion drawn from the evidence that the relationship is causal. Examination of the evidence might reveal that the relationship cannot be established, and thus a health claim cannot be based on the relationship.

Note that food businesses wishing to seek pre-approval for a food-health relationship underpinning either a high level or general level health claim (for inclusion in Standard 1.2.7) must follow the requirements for an application to FSANZ given in the *Application Handbook* (<http://www.foodstandards.gov.au/code/changes/pages/applicationshandbook.aspx>).

High level health claims must be based on a food-health relationship that has been pre-approved by FSANZ and included in section S4—4.

1.2 Requirements for notifying FSANZ of an established food-health relationship

Standard 1.2.7 requires that, if a general level health claim is based on a food-health relationship that has been established through a systematic review, the person making the claim must notify FSANZ of the food-health relationship and certify that the relationship was established by a process of systematic review as described in Schedule 6.

¹ Refer to Standard 1.1.2 (section 1.1.2—2) for the definition of general level health claim and other relevant terms. The Code is available via the FSANZ website (<http://www.foodstandards.gov.au/code/Pages/default.aspx>).

Further information about the notification process can be found at <http://www.foodstandards.gov.au/industry/labelling/fhr/Pages/notifications.aspx>. Note that food businesses are not required to provide FSANZ with the systematic review.

Food-health relationships notified to FSANZ are publicly listed on the FSANZ website.

1.3 Provision of records to a relevant authority

Paragraph 1.2.7—19(1)(d) requires the person making the claim, on request by a relevant authority (i.e. an enforcement agency), to provide records that demonstrate the systematic review was conducted in accordance with Schedule 6 and that the notified relationship is a reasonable conclusion of the systematic review.

When preparing a ‘systematic review report’ as described in this document, you should consider the need to be able to provide a ‘record’ that meets the requirements of paragraph 1.2.7—19(1)(d), if asked by regulatory authorities.

2 Overview of substantiating a food-health relationship

2.1 Substantiation

Schedule 6 allows two approaches for establishing a food-health relationship via a process of systematic review. One approach involves undertaking a systematic review by reviewing the original (also called primary) literature. The second approach allows a food business to use an existing systematic review and update it as described in paragraph S6—2(h) (see section 3.8 below). Both methods involve critical appraisal and quality assessment of the evidence.

‘Systematic reviews’ consist of a clearly formulated question and use of systematic and explicit methods to identify, select, critically appraise, and extract and analyse data from relevant research (Green and Higgins 2009). For the substantiation of food-health relationships, the relevant evidence must include studies in humans using experimental or observational designs. Figure 1 (page 5) shows that there are a number of points in the process of a systematic review where a food business might decide that the currently available information indicates the review is or is not worth pursuing any further at this time.

Establishing a food-health relationship using a process of systematic review is guided by the following principles (Health Canada 2011):

- **Systematic Approach:** a methodical, consistent approach to examining the relevant studies.
- **Transparency:** literature search strategies, selection and evaluation are fully disclosed and can be replicated.
- **Comprehensiveness:** all relevant evidence pertaining to the food-health relationship is captured, including evidence in favour and not in favour of the food-health relationship.
- **Evidence in humans:** a food-health relationship cannot be established from animal and *in vitro* studies alone. Studies in humans are essential.
- **Causality:** demonstration of causality is based on the quality and quantity of direct evidence which investigates the food-health relationship. Indirect or mechanistic evidence is not, sufficient by itself.

Reports from a number of government departments and non-government organisations often now contain systematic reviews that might be suitable for updating, depending on the

food-health relationship of interest to a food business. Peer-reviewed literature is another source of existing systematic reviews.

Many national and other bodies produce 'recommendations' or 'guidelines'. These recommendations/guidelines are often based on multiple sources of information. For example, the committee preparing a recommendation/guideline might have conducted reviews (systematic or otherwise) investigating the relationship of a number of outcomes to a single food or property of food. Consumer research or cost-benefit analysis might also have been conducted and included in the decision about whether to make a recommendation or not, and how to express it. Consequently, a recommendation/guideline is often based on considering a wider range of issues than a systematic review of a single food-health relationship, which is more narrowly focused.

It is possible that the work from which recommendations/guidelines are derived includes a systematic assessment of single focus relationships that might be useful for food businesses to examine further for relevance and to update. However, the recommendations/guidelines themselves might be expressed as dietary information and might not include a health effect. A health effect is an essential component of a food-health relationship. In closing, it should not be assumed that recommendations/guidelines are based solely on a systematic review of a single relationship, or meet the criteria to assess a food-health relationship for health claims purposes.

2.2 Preparation of the systematic review report

Scientific best practice suggests that a systematic review should be prepared by people with appropriate skills and qualifications for appraisal of data arising from clinical trials and epidemiological studies. It may also be useful to have a systematic review report peer reviewed. Both the author and reviewer would be expected to have a tertiary degree (of at least three years duration) in a scientific or health-related discipline and one or more of the following:

- training in critical appraisal or biostatistics from a tertiary institution
- a post-graduate degree (eg. MSc or PhD) in a scientific or health related discipline
- a specialist medical or health qualification.

2.3 Structure of the systematic review report

The structure of the report of the systematic review is not prescribed in Schedule 6. Food businesses might find the following suggested structure useful:

- executive summary
- qualifications of the author(s) and peer reviewer(s) if used
- description of the food/property of food
- description of the health effect
- food-health relationship examined, including direction (e.g. increase, decrease, maintenance etc) of effect and, if relevant, target population
- description of literature search strategy; inclusion and exclusion criteria, search terms, which databases were searched and how any unpublished evidence was ascertained
- summary of key information from selected studies in tabular form (as required by paragraph S6—2(d))
- assessment of the quality of included studies, including description of the quality assessment method
- assessment of the consistency of the association and demonstration of causality
- a conclusion about whether a causal relationship has been established

- assessment of the effective amount of the food or property of food and whether it can be consumed by the target population (if relevant) or the whole population in the Australian and New Zealand dietary context
- reference list or final list of studies (including unpublished studies)
- appendices.

If the approach chosen for substantiating a food-health relationship is by updating an existing systematic review, then scientific best practice suggests that the systematic review report include the following information in addition to the information above:

- an assessment of the quality of the existing systematic review, based on the requirements of Schedule 6
- a full copy of the existing systematic review
- a description of how and when the existing systematic review was updated.

In terms of content and structure, Schedule 6 requires that, among other things, the updated existing systematic review:

- be relevant (that is, demonstrate that the food-health relationship described in the existing and updated systematic review is based on the same, or is within the scope of, the proposed food-health relationship)
- includes all relevant data
- provides the information and demonstrates the conclusions required by paragraph S6—2(g), including whether or not a causal relationship has been established between the food (or property of food) and the health effect.

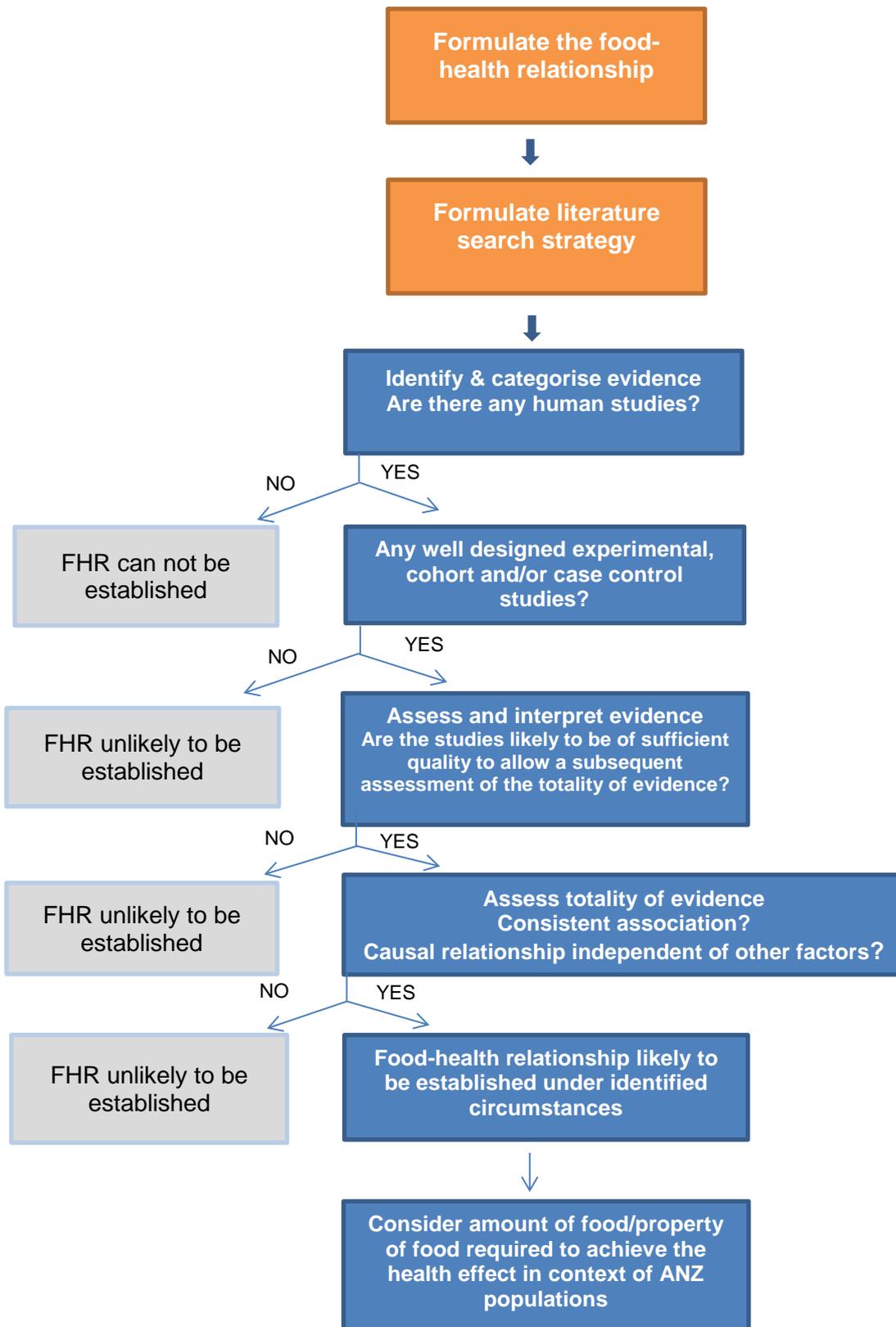
3 The systematic review

3.1 Reference material for conducting a systematic review

There are a number of handbooks, textbooks and papers about how to conduct a systematic review and appraise evidence, including:

- Health Canada (2009) Guidance document for preparing a submission for food health claims. 2009. http://www.hc-sc.gc.ca/fn-an/legislation/guide-ld/health-claims_guidance-orientation_allegations-sante-eng.php (accessed 15 March 2013)
- Health Canada (2011) Guidance Document for preparing a submission for food health claims using an existing systematic review. <http://www.hc-sc.gc.ca/fn-an/legislation/guide-ld/systemat-revi-sub-eng.php> (accessed 25 March 2013)
- GRADE Series. <http://www.jclinepi.com/content/jce-GRADE-Series> (accessed 15 March 2013)
- Green S, Higgins JPT (2011) Cochrane handbook for systematic reviews of interventions. Version 5.1.0. <http://handbook.cochrane.org/> (accessed 15 March 2013)
- NHMRC (National Health and Medical Research Council) (1999) How to review the evidence: systematic identification and review of the scientific literature. www.nhmrc.gov.au/files_nhmrc/publications/attachments/cp65.pdf (accessed 15 March 2013).

Figure 1: Decision points in the systematic review process of a food-health relationship (FHR), based on reviewing the original literature



Of these, two focus specifically on the health claims context (Health Canada 2009, 2011). Some, such as the Cochrane Handbook (Green and Higgins 2011) and certain NHMRC reports (NHMRC 1999) also focus on the appraisal of a single relationship. By contrast, others consider individual relationships but also include an appraisal of a guideline or recommendation that might be derived by simultaneously considering one or more relationships and other information such as cost-benefit analysis together (for example, the GRADE system). For the latter type, only that part of the reference that relates to individual relationships would be relevant to the health claims context. The terminology is not consistent in this field and users are advised to examine the concepts described in the texts and not assume that terminology in the texts is consistent with the terminology used in Standard 1.2.7.

For sections 1.2.7—18, 1.2.7—19, and 1.2.7—20, a systematic review must include the eight elements set out in paragraphs S6—2(a) to S6—2(g) in all cases; and also paragraph S6—2(h) if updating an existing systemic review.

Suggestions on how to undertake these elements are discussed below.

3.2 Description of the food-health relationship

Schedule 6 Required elements of a systematic review

S6—2(a) A description of the food or property of food, the health effect and the proposed relationship between the food or property of food and the health effect.

3.2.1 Description of the food or property of food

One approach to describing the food or property of food² that is the subject of the proposed food-health relationship might be to state clearly whether it is a food group (e.g. vegetables or fruit), a single ingredient food (e.g. banana), a food with more than one ingredient (e.g. chewing gum, bread) or a property of food that is either inherent or added (e.g. a nutrient, an ingredient, a component of an ingredient, or other substance or ingredient of food). Such a classification may facilitate the examination of retrieved literature to determine whether or not each study has investigated the food or property of food that is of interest.

Further characterisation might be desirable. For example, for:

- a food group (e.g. 'fruit') – the range of foods included could be described
- a single ingredient food (e.g. 'banana') – the genus, species and variety could be specified
- a food with more than one ingredient (e.g. 'bread') – ingredients could be described
- a property of food that may either be added or inherent – the common or usual name, the source or specifications including the CAS (Chemical Abstract Service) number and/or patent could be given if appropriate.

If the property of food has been added to the food under a specific permission in Australia and New Zealand, it might be useful to note which standard in the Code permits its addition to food.

² 'property of food' is defined in Standard 1.1.2 as 'a component, ingredient, constituent or other feature of food'.

Other aspects of the food or property of food that might be relevant to describe could be:

- the method used to measure consumption of the food or property of food or the amount of a food or property of food in a food product
- the food matrix (e.g. 'dairy foods') or other important nutritional or ingredient characteristics
- production methods, storage conditions etc.

Depending on how a food-health relationship is considered, it may be useful to limit, in the first instance, the description of a property of food to a single matrix. For example, if a systematic review considered the relationship for the property of food when consumed as a supplement, then it would be important to consider whether evidence for this relationship, based only on supplement intake studies, still applies when that property is present in a food matrix. Establishing bioequivalence for a property of food in different food matrices may be considered later in the substantiation process, such as when assessing the relevance of the review. This aspect is described in section 3.7.3. In general, a property of food, having a technical specification and present in two or more different food matrices, can be said to be bioequivalent if their bioavailabilities (presence in plasma), after administration of the same molar quantity, are similar to such a degree that their efficacy can be expected to be essentially the same.

3.2.2 Description of the health effect

A wide range of parameters are captured under the term 'health effect' as defined in section 1.1.2—2). The definition of 'health effect' refers to what might be called 'the outcome' in some scientific manuscripts. For each food-health relationship a specific health effect should be described. One approach for describing the health effect is to state formally the parameter of interest. For some parameters, it might be relevant to state a measurement technique as part of the description.

Scientific best practice suggests that consideration is needed to determine how to review a body of literature that includes studies using different measurement methods. For example:

- Serum 25-hydroxy vitamin D can be measured using radioimmunoassay, high performance liquid chromatography or liquid chromatography-tandem mass spectrometry. Appreciable bias and variability between laboratories and between assays is sufficient to affect between-study comparisons. Those doing a systematic review of the effects vitamin D on a health effect would need to consider whether studies using any of these methods would be acceptable in the systematic review, or not.
- Growth in children can be compared to growth charts released by the UK, WHO and the USA. If a systematic review was using growth as an outcome, consideration would need to be given to whether the systematic review should be restricted to studies that used only one of these growth charts or not.
- Development in children or cognitive functioning in adults can be measured using a range of different tools, which are not necessarily comparable.

One approach for defining some types of health effect might be to use the International Classification of Disease (ICD) code (<http://www.who.int/classifications/icd/en/> (accessed 15 March 2013)). This might be more relevant for health effects which are conditions assessed in the health and medical systems. Other health effects may not be captured in ICD. The European Food Safety Authority (EFSA) has a series of guidance documents on

health effects that can be measured in different body systems that might be useful (<http://www.efsa.europa.eu/en/applications/nutrition>).

3.2.3 Description of the proposed food-health relationship

Paragraph S6—2(a) requires that a systematic review must include a description of, among other required elements, the proposed relationship between the food or property of food and the health effect.

The proposed direction of effect of the food or property of food on the health effect should be described. For example, the review should identify whether increasing intake of a food or property of a food may increase, decrease or maintain the level of the health effect. If the relationship relates only to a specific target population group (e.g. the elderly) then this should be stated.

There are a number of texts and handbooks on systematic reviews. Some of these might refer to the description of the food-health relationship as the research or review question.

3.3 Retrieval of scientific evidence – systematic review based on the original literature only

Schedule 6 Required elements of a systematic review

S6—2(b) A description of the search strategy used to capture the scientific evidence relevant to the proposed relationship between the food or property of food and the health effect, including the inclusion and exclusion criteria.

S6—2(c) A final list of studies based on the inclusion and exclusion criteria. Studies in humans are essential. A relationship between a food or property of food and the health effect cannot be established from animal and *in vitro* studies alone.

3.3.1 Development of the search strategy

Scientific best practice is to develop and document a relevant, comprehensive, systematic and reproducible search strategy to capture the totality of evidence from studies investigating the food-health relationship. This includes evidence that supports, or does not support, the relationship being investigated.

Schedule 6 (paragraph S6—2(c)) states that studies in humans are essential. However, subject to this requirement, animal and *in vitro* studies may provide useful supporting information such as information about the biological plausibility of a food-health relationship.

3.3.1.1 Inclusion and exclusion criteria

In addition to the information about the food or property of food, the health effect and the relationship, additional parameters may need to be defined to allow a relevant search to be performed. The mnemonics PICO (ie. population, intervention, comparator, outcome; Russell et al. 2009), PICOT (Riva et al. 2012) or PECOT (Heath 2009, Davies, 2011) are often used to summarise the information that can be considered when developing a question for doing a systematic review. Some mnemonics include an S for study design (e.g. PICOTS). These mnemonics contain some of the key items from the description of the food-health relationship but the expansion might be useful for identifying the relevant literature as follows:

- P: population – does the relationship apply to the whole population or a subset?
- I or E: what is the intervention (in a trial) or exposure (in an observational study) (i.e. the food or property of food)?
- C: what is the comparison or comparator? (e.g. A placebo? Comparison of high versus low intake? Isoenergetic replacement of carbohydrate with polyunsaturated fat?)
- O: what is the outcome (i.e. health effect)? It may be useful to include measurement method when specifying the outcome.
- T: timeframe (e.g. in studies lasting at least 3 months or 1 year etc.)
- S: study design (e.g. randomised controlled trial, cohort study, case-control study).

These, and similar parameters are often used to determine inclusion/exclusion criteria and to define the research question. Some types of parameters could be described as either inclusion or exclusion criteria for the search strategy. For example, if only studies on adults are to be retained for a review, then this could be described equally well as ‘adults 18 years and older are included’ or ‘children aged less than 18 years are excluded’. Other types of criteria are more appropriately described in a particular way. For example ‘studies of people with previously diagnosed cancer are excluded’. Sometimes double-barrelled terminology might be used, for example ‘folate including folic acid and 5-methyltetrahydrofolate but excluding folinic acid’. Some authors consider that ‘inclusion criteria’ should only refer to PICO(TS) items and ‘exclusion criteria’ should only refer to other items used to screen the literature.

Inclusion/exclusion criteria are used to search the electronic databases and then to filter the results. Some inclusion/exclusion criteria can be implemented at the search stage, for example the literature is typically classified by study design, age ranges or sex in the electronic databases. Others, such as a criterion about duration of a study or use of a specific measurement tool, might require the abstract, and possibly the entire paper, to be read before a decision can be made about whether the paper meets the inclusion/exclusion criteria.

It can be difficult to decide whether to restrict a search using the inclusion/exclusion criteria or to conduct a wider search and then justify extrapolation (generalisation) of the results to a particular group. For example, if a health claim is planned that would target adults, then the stated food-health relationship might specify this. In this case, studies in children identified in the literature would be discarded. There may be other instances when the target population group would not be considered until later in the substantiation process. For example, if the focus is on women, but studies have been conducted primarily in men, then studies in men might be included but formal consideration would need to be given as to whether the results can be extrapolated to women.

The following types of, and information about, inclusion/exclusion criteria are typically described:

- age, sex, possibly race/ethnicity
- exclusions related to pre-existing diseases or physiological conditions

- the terms and subject headings (e.g. Medical Subject Headings (MeSH)³ used; whether terms were 'exploded'
- criteria related to the intervention/exposure (i.e. the food or property of food), the comparator and the outcome (i.e. health effect).

Scientific best practice suggests that the decision not to use any potentially relevant search terms should be justified.

A copy of the search strategy/ies is often reported in systematic reviews. The systematic reviews found in the Cochrane Library (<http://www.thecochranelibrary.com/view/0/index.html> (accessed 12 April 2013)) contain examples of this type of documentation. The assistance of a librarian with experience in the health field may be helpful in developing a relevant and comprehensive search strategy.

3.3.1.2 Identifying the databases to search

Three important inclusion/exclusion criteria are the database(s) searched, criteria concerning language and the time period searched. Scientific best practice suggests searching at least two different databases. The following data sources are not considered to be suitable for inclusion in a systematic review for a food-health relationship:

- articles published in newspapers, magazines, newsletters, etc.
- books or book chapters for consumers or the general public
- information intended for the general public on the internet, such as Wikipedia.

How far back in time the search might extend would depend on the food-health relationship being investigated. For some topics, the property of food might have come to scientific attention only recently and the date of the first publication about it might be known. Similarly, if use of a specific measurement method is an inclusion/exclusion criterion, then the development date of the method might be ascertainable and could be used to set a boundary on the time period of the search. In other cases, a much longer time period might need to be searched. The inclusion or exclusion of non-English language literature needs to be considered.

The following types of inclusion/exclusion criteria relating to sources are typically described:

- the electronic databases searched (eg. Medline, CINAHL (Cumulative Index to Nursing and Allied Health Literature, Cochrane CENTRAL, Embase, PsycINFO)
- languages excluded from the search (if any)
- time period searched and reasons for choosing the time period
- any manual (non-electronic) search techniques employed, including hand-searching and the strategy used to identify any unpublished studies (see below).

3.3.1.3 Unpublished or proprietary material

The unpublished results of studies, including proprietary studies, can contribute to the evidence base for a food-health relationship provided they meet the same inclusion criteria as the published studies and that there is a systematic and documented approach to identifying all unpublished studies. Such studies may be identified in a variety of ways including formal hand-searching of books of abstracts from relevant conferences and

³ The National Library of Medicine's Medical Subject Headings (MeSH) is the controlled vocabulary used for indexing articles for the Medline subset of the PubMed database. MeSH terminology provides a consistent way to retrieve information where several different terms may be used for the same concept.

contacting authors of completed trials identified via trials registries. Examples of trial registries are:

- the Australian New Zealand Clinical Trials Registry
[http://www.ctc.usyd.edu.au/prospective-trial-registration/australian-new-zealand-clinical-trials-registry-\(anzctr\).aspx](http://www.ctc.usyd.edu.au/prospective-trial-registration/australian-new-zealand-clinical-trials-registry-(anzctr).aspx) (accessed 15 March 2013)
- the International Clinical Trial Registry Platform managed by WHO
<http://www.who.int/ictcp/en/> (accessed 15 March 2013)
- the registry managed by the US National Institutes of Health
<http://www.clinicaltrials.gov/> (accessed 15 March 2013).

It may be more difficult to find proprietary studies that have not been publicly released. Unpublished trials that meet the inclusion criteria, even those which were 'unsuccessful' and not written up extensively, should be included. This helps to ensure that the overall results of a systematic review are unbiased.

3.3.2 Filtering the retrieved evidence using inclusion/exclusion criteria

Searching electronic databases generally retrieves a larger number of studies than are relevant to a specific question. One way to screen studies is to firstly assess study titles, then abstracts, followed by the full text against the inclusion/exclusion criteria. For example if an inclusion criterion is that the study population has to be relevant to the Australian and New Zealand dietary context (for example in a population with similar food consumption or nutrient intake patterns), then it may be possible to exclude a study about vitamin A deficiency in Africa based on the title of the study alone. An aspect of the study design such as study duration may be able to be assessed from the abstract, and studies excluded, if they were conducted for less than a stated minimum period. Other details may only be able to be assessed from reading the full paper such as details of measurement methods used. The systematic reviews found in the Cochrane Library (<http://www.thecochranelibrary.com/view/0/index.html> (accessed 12 April 2013)) contain examples of this type of documentation.

Scientific best practice suggests documenting the number of studies at each stage of filtering in a flow diagram with a summary of reasons for exclusions. This type of diagram is called a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (<http://www.prisma-statement.org/Default.aspx> (accessed January 2016)). Appendix 1 shows an outline of a PRISMA flow diagram.

Further details about literature searching and filtering can be found in the Cochrane Handbook (Green and Higgins 2011) and a number of textbooks on systematic reviews. PubMed (<http://www.ncbi.nlm.nih.gov/pubmed> (accessed 12 April 2013)) has online tutorials about searching the Medline database using search terms.

Software is available which can help with identifying duplicates resulting from searching more than one database and managing the search and filtering processes.

Schedule 6 (paragraphs S6—2(b) and S6—2(c)) requires that the inclusion and exclusion criteria used to filter the evidence are described along with the final list of studies.

3.4 Tabulation of data from the final list of included studies

Schedule 6 Required elements of a systematic review

S6—2(d) A table with key information from each included study. This must include information on:

- (i) the study reference; and
- (ii) the study design; and
- (iii) the objectives; and
- (iv) the sample size in the study groups and loss to follow-up or non-response; and
- (v) the participant characteristics; and
- (vi) the method used to measure the food or property of food including amount consumed; and
- (vii) confounders measured; and
- (viii) the method used to measure the health effect; and
- (ix) the study results, including effect size and statistical significance; and
- (x) any adverse effects.

The aim of tabulating the included studies is to summarise the critical features of each study in a standardised and objective manner. This allows readers to quickly identify key aspects of each study and it also assists with assessing the quality of each study (section 3.5) and forming a judgement about the body of evidence (section 3.6). A table must include study features listed in subparagraphs S6—2(d)(i)–(x) for each included study but the order of presentation of the listed features is not prescribed.

There are a number of templates that could be used to summarise the features of each study. It should be noted that a template may need to be customised in order to reflect the requirements of paragraph S6—2(d).

Examples of templates can be found in systematic reviews produced by the Cochrane Collaboration (<http://www.thecochranelibrary.com/view/0/index.html> (accessed 12 April 2013)), other systematic reviews in the scientific literature and Appendix 2. It may be useful to group studies in the table according to study design (e.g. randomised controlled trial, cohort study (including nested case-control or case-cohort studies) or case-control study), by type of measurement method used or by some other relevant feature of the studies.

Unpublished material from whatever source needs to be included in tabulations and quality assessment. To avoid bias, scientific best practice suggests only including each study once in a systematic review. This means that any conference abstracts identified that were subsequently published as full papers would be excluded because they duplicate the study results. Similarly, if a cohort study has reported results for several waves of follow-up, then resulting data from only one wave of follow-up would be included in the systematic review for the same reason. However, it is possible that several papers might need to be read from large studies to obtain all the information needed to populate a table for one set of follow-up data.

3.5 Assessment of study quality

Schedule 6 Required elements of a systematic review

S6—2(e) An assessment of the quality of each included study based on consideration of, as a minimum:

- (i) a clearly stated hypothesis; and
- (ii) minimisation of bias; and
- (iii) adequate control for confounding; and
- (iv) the study participants' background diets and other relevant lifestyle factors; and
- (v) study duration and follow-up adequate to demonstrate the health effect; and
- (vi) the statistical power to test the hypothesis.

Paragraph S6—2(e) lists the study characteristics (as a minimum) that must be considered in a quality assessment of each included study. The aim of an assessment of the quality of each included study is to identify studies that are more likely to report unbiased results and therefore are of higher quality. One possible strategy is to have two independent reviewers provide ratings to appraise the quality of each study. If their assessments differ, the sources of differences can be identified and resolved through discussion.

3.5.1 Quality appraisal tools

There are a number of different tools that could be used to assess the quality of each included study although they cover a common core of concepts for assessing quality. These tools include:

- Health Canada 2009
- GRADE Working Group (<http://www.gradeworkinggroup.org/> (accessed 23 May 2013))
- National Collaborating Centre for Methods and Tools 2008
- Scottish Intercollegiate Guidelines Network (SIGN) 2004
- Green and Higgins 2011
- NHMRC 1999.

Note that the tool described by Health Canada (2009) is specifically applied to the health claims context (refer to Appendix 3 for some of the quality tables (reproduced with permission)). Other tools consider risk factor-disease outcome type relationships in general. Consistent with scientific best practice, the quality assessment tool and its rating system should be stated in the systematic review report. It should be noted that the available tools do not necessarily cover all the criteria specified in paragraph S6—2(e) and so elements of more than one system may be needed. In addition different aspects of quality are more or less important in different types of studies. For example, confounding is controlled by randomisation (if it is done well) in a randomised controlled trial whereas it needs to be dealt with by statistical adjustment in an observational study. Consequently a review that includes observational studies might give greater detail about how potential confounding was managed in the included studies than a review that included only well conducted randomised controlled trials. Background diet may not be an important consideration in a review where all trials were conducted in populations with dietary intakes similar to those in Australia and New Zealand. However, background diet might be important when interpreting, for example, the

results of a trial of omega-3 fatty acids conducted in a country with a substantially higher fish intake than Australia and New Zealand.

Readers need to be aware of the variable terminology among the texts and handbooks in this area. For example, the GRADE tool (Balshem et al. 2011), uses the term 'quality' to refer not only to an assessment of individual studies, but also to an assessment of a systematic review based on a collection of studies. Balshem et al. (2011) then extend the idea of quality to a recommendation that might be derived from multiple sources of evidence including a cost-benefit analysis. This concept is clearly different from the health claims context in which the term 'quality' refers only to an assessment of individual studies for a specific food-health relationship.

3.6 Assessment of the body of evidence and conclusion

Schedule 6	
Required elements of a systematic review	
S6—2(f)	An assessment of the results of the studies as a group by considering whether: <ul style="list-style-type: none">(i) there is a consistent association between the food or property of food and the health effect across all high quality studies; and(ii) there is a causal association between the consumption of the food or property of food and the health effect that is independent of other factors (with most weight given to well-designed experimental studies in humans); and(iii) the proposed relationship between the food or property of food and the health effect is biologically plausible; and
S6—2(g)	A conclusion based on the results of the studies that includes: <ul style="list-style-type: none">(i) whether a causal relationship has been established between the food or property of food and the health effect based on the totality and weight of evidence; and

3.6.1 Assessment of the body of evidence

Following the assessment of the quality of each study, the findings across all studies are considered together. The totality of evidence should be considered which means all studies on a topic that meet the pre-determined criteria are included even if they have results that do not support establishment of a food-health relationship, i.e. studies with equivocal, opposing or null effects. The findings are examined to determine whether there is a consistent association. Subparagraph S6—2(f)(i) also requires that consideration be given to whether there is a consistent association between the food or property of food and the health effect across all high quality studies.

There are several tools available for combining studies and examining consistency of results. Formal quantitative meta-analysis is one tool. Health Canada provides a tool that does not involve meta-analysis (http://www.hc-sc.gc.ca/fn-an/legislation/guide-ld/health-claims_guidance-orientation_allegations-sante-eng.php (accessed 23 May 2013)). Appendix 4 shows the GRADE system ratings of quality of evidence (Balshem et al. 2011) and what characteristics affect the rating given to the overall relationship. In addition to numerical combination of study results using techniques such as meta-analysis, a qualitative assessment is used to help determine whether an association should be regarded as causal. Criteria for this part of the assessment include, but are not limited to:

- strength of association (size of the effect)
- dose-response relationship
- reversibility or sustainability of the health effect
- consistency of results when studies are done by different authors or in different samples
- comparability of methods used to measure intake or health effects in the study design
- temporality (intake of the food or property of the food precedes the health effect)
- independence of association (other possible explanations have been ruled out)
- biological plausibility
- whether additional well conducted (high quality) studies could potentially alter the association seen across the high quality studies.

Causality also includes the idea that there is no doubt that intake of the food/property of food occurred before the health effect. This is often referred to as the 'temporal' assumption (Hill 1965). This is a different idea from consistency. A relationship could be consistent across studies, for example, cross-sectional studies might consistently report an association between high LDL-cholesterol levels and high consumption of polyunsaturated margarine. However this study design would not allow a determination of whether the elevated cholesterol levels occurred before or after margarine consumption commenced. One possible explanation is that people with elevated cholesterol were following medical advice to increase polyunsaturated fat intake. Scientific best practice for assessing whether a relationship might be causal typically excludes cross-sectional, ecological studies and case-series because it is difficult to identify the temporal direction. Case-control studies are often, but not always, excluded owing to the uncertainty in whether the case subjects are able to recall their dietary intake before the onset of the disease. Of the various designs, trials and prospective cohort studies generally provide the most certainty about the criterion of temporality.

Another important criterion for determining that a causal relationship between the food or property of food and the health effect exists (see Schedule 6, subparagraph S6—2(f)(ii)) is that it is independent of other factors, usually referred to as confounders. In a randomised controlled trial, other explanations are eliminated by good randomisation methods, including masked allocation and blind assessment of outcomes. Statistical control of confounding might also be necessary in a trial. In observational studies, all important known factors (confounders) would have been measured and then controlled for statistically. Scientific best practice suggests checking that known confounders (which includes other causes of the health effect) have been measured, because if they have not been measured, then they cannot be controlled statistically. Even then it can be hard to be certain that confounders have been measured adequately. When doing a quality appraisal, poorly conducted trials would have their quality rating downgraded and there are certain features that would allow some observational studies to have their quality rating upgraded (Appendix 4; Balshem et al. 2011). Thus the quality assessment, and not simply the study design, affects the final assessment of the extent to which the body of evidence allows a causal conclusion to be drawn.

The strength of the association (size of the relative risk) is another important criterion. There are rare instances when the relative risk from a set of well-designed observational studies is so large (e.g. 10 or 20) that a causal conclusion can be drawn. These large relative risks may be found for diseases that have a single cause. However, in the field of modern nutrition, the health effects of interest have multiple causes and so the relationships are typically small and have considerable confounding that needs to be removed if observational data are used.

It is not possible to predict the number of studies that would be needed to allow a causal relationship to be established following the finding of a consistent association and consideration of the other criteria. This is due to the variation in the magnitude of the association, study quality, sample size and control for confounding across the wide range of food-health relationships that might underpin health claims. One way of thinking about causality might be to consider whether it is likely or not that another large, well-conducted study would have such different results from the available studies that the conclusion from the systematic review would be altered importantly.

As noted above, animal and *in vitro* studies might be helpful in assessing whether the food-health relationship of interest is biologically plausible even though they cannot alone substantiate a food-health relationship.

3.6.2 Conclusion

After considering all the relevant data, all the criteria noted above, which studies are higher quality and should be given greater weight, and the consistency of the association in the high quality studies, a conclusion should be made about whether or not a causal relationship is established. If a food-health relationship is causal rather than simply being an association, then additional high quality studies reported in the future would support rather than overturn the conclusion that the relationship is causal. Thinking about the amount of high quality data that would be required to alter the conclusion might be one way to consider if an association is robust enough to be regarded as causal. Not all relationships that show a consistent association, even in high quality studies, would be based on enough data to be regarded as causal.

In addition to the Code, health claims are also subject to food and consumer laws in Australia and New Zealand. Food businesses need to determine how often they should scan the evidence to determine if new studies have been conducted that would alter the conclusions of food-health relationships that underpin general level health claims, to ensure that claims are in compliance with these laws.

3.7 Applicability to Australia and New Zealand

Schedule 6	
Required elements of a systematic review	
S6—2(f)	An assessment of the results of the studies as a group by considering whether: <ul style="list-style-type: none"> (iv) the amount of the food or property of food to achieve the health effect can be consumed as part of a normal diet of the Australian and New Zealand populations.
S6—2(g)	A conclusion based on the results of the studies that includes: <ul style="list-style-type: none"> (ii) where there is a causal relationship between the food or property of food and the health effect: <ul style="list-style-type: none"> (A) the amount of the food or property of food required to achieve the health effect; and (B) whether the amount of the food or property of food to achieve the health effect is likely to be consumed in the diet of the Australian and New Zealand populations or by the target population group, where relevant.

3.7.1 Consumption of the food or property of food and consideration of the amount to achieve the health effect

Subparagraphs S6—2(f)(iv) and S6—2(g)(ii) of Schedule 6 require that the amount of the food or property of food needed to achieve the health effect is considered along with the amount of the food or property of food likely to be consumed in the diet of the Australian and New Zealand populations or by the target population.

A representative survey of the population of interest is regarded as providing the best information about intake (NHMRC 1999). When the whole population, or an age-sex subgroup of the general population, is the target, a representative national or state survey would provide an appropriate source of information. If the claim is directed at a subgroup for which representative surveys do not give a reliable estimate, then a well-sampled special-purpose survey would be desirable.

Scientific best practice provides several methods for deriving 'usual' intake estimates from survey data. Assessing usual intake would be relevant if the claim derived from the food-health relationship relates to the intake of the food or property of food over a long period of time.

Consideration should be given to what foods can carry the claim, taking into account the amount of the food or property of food required to achieve the health effect and the distribution of the food or property of food in the food supply. For example, the amount of the property of food required for the health effect may not necessarily have to be present in a single serving of the food, particularly if that property of food is widespread in the food supply.

3.7.2 The target population group

As noted above, a food-health relationship underpinning a claim might focus on a specific target population group rather than the general population. There are several approaches to assessing the food-health relationship for a target population group. In some cases, the systematic review can be restricted to include only studies on the target population group. In other cases, studies with a different or wider target population group would be included in the systematic review, and then consideration would be given to whether the results of the systematic review could be extrapolated (generalised) to the target population group of interest. For example, the literature might contain only studies on subjects with a specific condition (e.g. overweight) and so consideration would be given to whether the relationship could be extrapolated to target population groups with different weight status, given that the relationship had been established from studies of overweight people.

3.7.3 Extrapolation from supplements or other matrices

It may be necessary to include a formal assessment of whether an established relationship based on supplements can be extrapolated to food matrices. Sometimes, a 'bridging' study which compares bioavailability of a property of food in a range of food matrices can be used to determine if a supplement intake relationship is applicable.

If studies in the systematic review have been conducted in one food matrix, then consideration needs to be given to the basis for extrapolating the relationship to apply to a claim on a different food matrix (e.g. phytosterols in edible oil spreads versus low fat cheese). Similarly, if the systematic review contained studies of one food (e.g. oats), then consideration would need to be given to the rationale for extrapolating the relationship if a claim were intended to be used on other foods in the same food group (e.g. wheat) or beyond the food group (e.g. legumes).

3.8 Updating an existing systematic review

Schedule 6 Required elements of a systematic review

- S6—8 An existing systematic review may be used if it is updated to include:
- (a) the required elements 1 to 6 above for any relevant scientific data not included in the existing systematic review.
 - (b) the required element 7 above incorporating the new relevant scientific data with the conclusions of the existing systematic review.

3.8.1 Selection of existing systematic review

Food businesses can update and rely on an existing systematic review in certain circumstances. An existing systematic review, if updated, must meet the requirements for establishing a food-health relationship as described in Schedule 6.

In choosing an existing systematic review as a starting point, scientific best practice would suggest that the food-health relationship examined in the existing systematic review closely aligns with or be identical to the food-health relationship of interest to the food business (refer to section 3.2).

There are a number of systematic reviews of food-health relationships which appear as individual papers in the peer reviewed literature or as part of larger sets produced by government bodies such as the National Health and Medical Research Council (NHMRC), Health Canada or the World Health Organization. Non-government organisations such as the Cochrane Collaboration also produce systematic reviews.

Some groups have started to articulate the degree of certainty that they have in the food-health relationship they have examined in their systematic reviews. These self-ratings for relationships by previous authors might provide useful, although not definitive, assistance to food businesses in determining the utility of investigating certain topics. Rating systems which might be seen in the literature include the NHMRC system which rates relationships from A-D (Merlin et al. 2009) and the GRADE system, used by the Cochrane Collaboration and the World Health Organization, which rates relationships from ++++ to + (<http://www.iclinepi.com/content/jce-GRADE-Series> (accessed 15 March 2013)).

3.8.2 Updating an existing systematic review against subparagraphs S6—2(a) to S6—2(g)

Subparagraph S6—2(h) provides that an existing systematic review may be used if it is updated in accordance with that paragraph.

Scientific best-practice suggests there are two main aspects to such an update. Firstly, an existing systematic review should be assessed to determine whether it captured all relevant data within the time period that the authors searched. Secondly, any relevant data published outside the stated time period of the existing systematic review should be included in the updated systematic review. In recent years, many journals have started to release papers intended for future publication on their websites prior to print publication. These papers are often picked up and included in electronic databases prior to print publication. Therefore papers which are printed early in the update period might have already been included in the existing review. This should be checked to ensure that some studies are not included in the review twice.

As the update needs to combine any additional relevant scientific data identified, included and assessed using paragraphs S6—2(a) to S6—2(f) with scientific data in the existing systematic review to draw a conclusion, one approach would be to assess whether the methods used in the existing systematic review followed paragraphs S6—2(a) to S6—2(f) closely. Part of this assessment might include considering the validity of the conclusions presented by the authors of the existing systematic review in light of the data shown in the existing systematic review. If the authors of the existing systematic review concluded that a causal relationship had been demonstrated, then the food business would need to consider whether this determination was supported. Alternatively, some existing systematic reviews might have found an association but not drawn a causal determination. It might be the purpose of the update to examine whether new data would allow a causal relationship to be established or not.

One approach to check for additional relevant data would be to run the search strategy described by the authors of the existing systematic review in the same electronic databases (and hand searching the same series of conference abstracts if relevant). If the authors of the existing systematic review have clearly stated the last date included in their search, then the updating search can start at that point. For example, if an existing systematic review searched literature published between 1 January 1990 and 31 December 2009, then the updated search could start from 1 January 2010. There is a need to consider whether the start date of a literature search had missed relevant information published before that date; for example owing to a delay between publication and inclusion in the electronic database. The use of unpublished reports in this context would be the same as that covered in section 3.3 above. The updated review should use the same, or similar, search criteria.

The amount of additional relevant literature on a topic depends on how recent the existing systematic review is, the quality of the original review (including how thorough the authors of the existing review were) and also on the amount of new research in the area. It is possible that no additional relevant literature would be identified in searches seeking to update some food-health relationships whereas there might be a large quantity of additional relevant literature for other food-health relationships.

Any additional relevant scientific data should be filtered and assessed as described above for a systematic review that includes only original literature (refer to sections 3.3.2, 3.4 and 3.5).

Subparagraph S6—2(h)(ii) indicates that any additional relevant data need to be considered together with data already in the existing systematic review to make an assessment about consistency of association across high quality studies, etc., to derive a conclusion about whether the total body of evidence (data in the existing systematic review combined with any additional relevant data), indicates a causal association (refer to section 3.6). Note that the amount of the food or property of food required to achieve the health effect and whether this can be consumed as part of a normal diet in Australia and New Zealand also need to be considered when updating an existing systematic review (refer to section 3.7).

References

Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gotzsche PC, Lang T. (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 134: 663-694.

Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH. (2011) GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 64: 401-406. [http://www.jclinepi.com/article/S0895-4356\(10\)00332-X/fulltext](http://www.jclinepi.com/article/S0895-4356(10)00332-X/fulltext) Accessed 12 April 2013

Codex Alimentarius Commission. (1997) Guidelines for use of nutrition and health claims CAC/GL 23-1997.

www.codexalimentarius.org/input/download/standards/351/CXG_023e.pdf Accessed 25 March 2013

Davies KS. Formulating the Evidence Based Practice Question: A Review of the Frameworks Evidence Based Library and Information Practice 2011, 6.2 Available at <https://ejournals.library.ualberta.ca/index.php/EBLIP/article/viewFile/9741/8144> (Accessed 4 May, 2016)

EFSA (European Food Safety Authority) (2011) Scientific and technical guidance for the preparation and presentation of an application for authorisation of a health claim (revision 1). <http://www.efsa.europa.eu/en/efsajournal/pub/2170.htm> Accessed 25 March 2013

FSANZ (Food Standards Australia New Zealand) (2008) Proposal P293 – Nutrition, Health and Related Claims. Final Assessment Report. http://www.foodstandards.gov.au/code/proposals/documents/P293%20Health%20Claims%20FAR%20and%20Att%201%20_%202%20FINAL.pdf Accessed 6 January 2016

GRADE Working Group <http://www.gradeworkinggroup.org/> Accessed 15 March 2013

Green S, Higgins JPT (2011) Cochrane handbook for systematic reviews of interventions, Version 5.1.0. <http://www.cochrane.org/training/cochrane-handbook> Accessed 25 March 2013

Grossklaus R (2009) Codex recommendations on the scientific basis of health claims. *Eur J Nutr*. 2009 Dec; 48 Suppl 1: S15-22. <http://link.springer.com/article/10.1007%2Fs00394-009-0077-z> Accessed 25 March 2013

Health Canada (2009) Guidance document for preparing a submission for food health claims. http://www.hc-sc.gc.ca/fn-an/legislation/guide-ld/health-claims_guidance-orientation_allegations-sante-eng.php Accessed 25 March 2013

Health Canada (2011) Guidance Document for preparing a submission for food health claims using an existing systematic review. <http://www.hc-sc.gc.ca/fn-an/legislation/guide-ld/systemat-revi-sub-eng.php>

Heath JM. Finding the evidence – Planning your search. In: Grimmer-Somers, K Kumar, S, Worley, A, Young, A. Practical tips in finding the evidence: An Allied Health primer. 1st ed. Manila: UST Publishing House; 2009.

Hill AB (1965) The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine* 58 (5): 295–300.

Merlin T, Weston A, Toohar R et al. NHMRC levels of evidence and grades for recommendations for developers of guidelines. Canberra: NHMRC, 2009. <http://www.nhmrc.gov.au/guidelines/resources-guideline-developers> Accessed 25 March, 2013

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(7): e1000097. doi:10.1371/journal.pmed.1000097 <http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1000097> Accessed 12 April 2013

National Collaborating Centre for Methods and Tools (2008) Quality Assessment Tool for Quantitative Studies. Hamilton, ON: McMaster University. (Updated 13 April 2010). Retrieved from <http://www.nccmt.ca/registry/view/eng/14.html> Accessed 8 May 2013

NHMRC (National Health and Medical Research Council) (1999) How to review the evidence: systematic identification and review of the scientific literature. www.nhmrc.gov.au/files_nhmrc/publications/attachments/cp65.pdf. Accessed 12 April 2013

Quality Assessment Tool for Quantitative Studies. Canada: Effective Public Health Project. <http://www.ephpp.ca/tools.html> Accessed 15 March 2013

Riva JJ, Malik KMP, Burnie SJ, Endicott AR, Busse JW (2012) What is your research question? An introduction to the PICOT format for clinicians *J Can Chiropr Assoc*. 2012 September; 56(3): 167–171. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3430448/> Accessed 23 May 2013

Russell R, Chung M, Balk EM, Atkinson S, Giovannucci EL, Ip S, Mayne ST, Raman G, Ross AC, Trikalinos T, West KP, Lau J (2009). Issues and Challenges in Conducting Systematic Reviews to Support Development of Nutrient Reference Values: Workshop Summary. Nutrition Research Series, Vol. 2. Technical Reviews, No. 17.2. <http://www.ncbi.nlm.nih.gov/books/NBK44081/> Accessed 17 April 2013

Scottish Intercollegiate Guidelines Network (SIGN) (2004) Annex C: Critical Appraisal – Notes and checklists. Methodology Checklist 2: Randomized Controlled Trials. Edinburgh: SIGN. <http://www.sign.ac.uk/methodology/checklists.html> Accessed 6 January 2016

Scottish Intercollegiate Guidelines Network (SIGN) (2004) Annex C: Critical Appraisal –Notes and checklists 3 & 4: Cohort and case-control studies. Edinburgh: SIGN. <http://www.sign.ac.uk/methodology/checklists.html> Accessed 6 January 2016

TGA (Therapeutic Goods Administration) (2012) Consultation: Evidence required to support indications for listed medicines (excluding sunscreens and disinfectants). <http://www.tga.gov.au/newsroom/consult-cm-evidence-listed-medicines-120827.htm> Accessed 25 March 2013

USFDA (United States Food and Drug Administration) (2009) Guidance for industry: Evidence-based review system for the scientific evaluation of health claims – Final. <http://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/LabelingNutrition/ucm073332.htm> Accessed 25 March 2013

Department of Health 2012, Victoria Population Health Survey 2010. State Government of Victoria, Melbourne. <http://www.health.vic.gov.au/healthstatus/survey/vphs.htm> Accessed 25 March 2013

Glossary

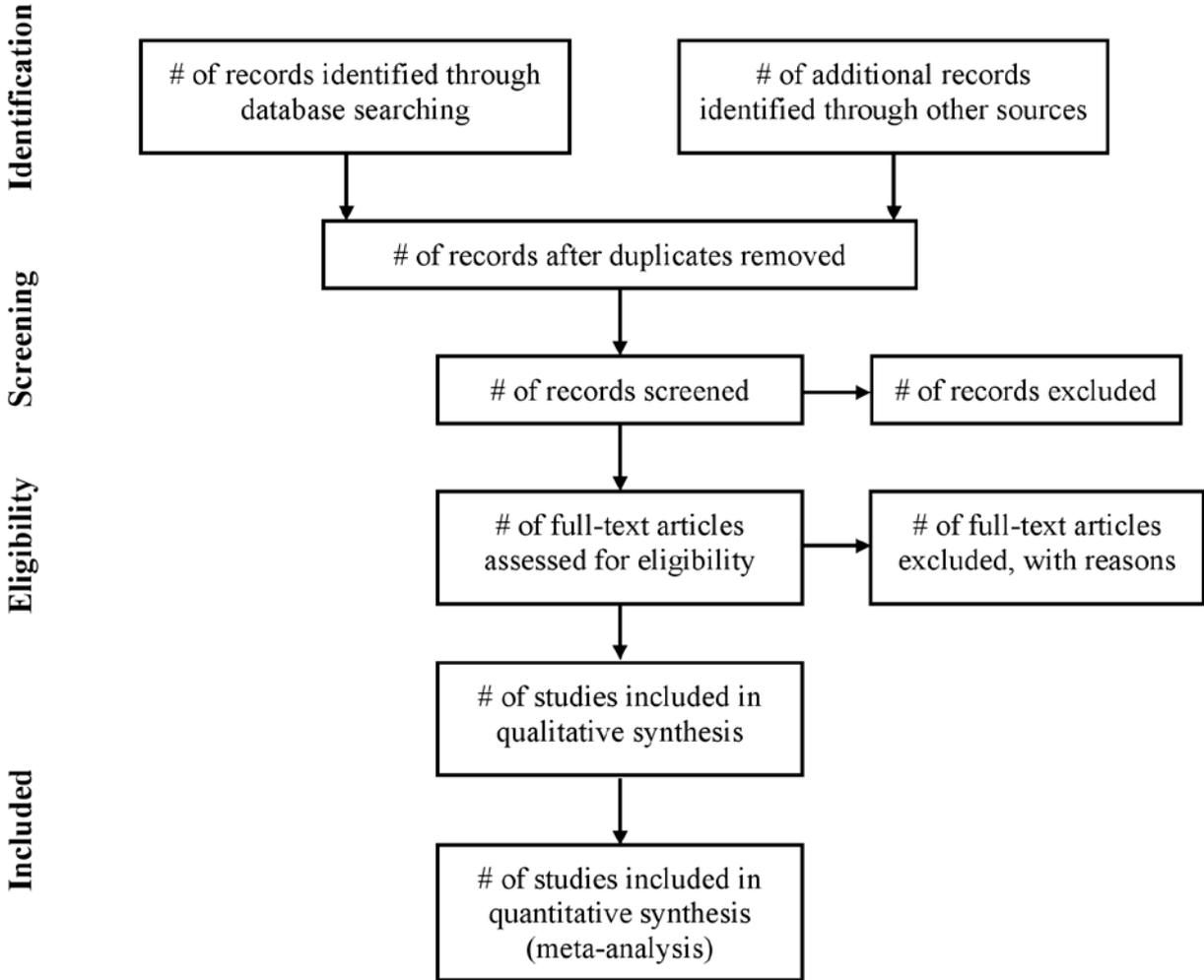
Bias	Systematic deviation of a measurement from the 'true' value leading to either an over- or underestimation of the treatment effect. Bias can originate from many different sources, including measurement, interpretation, publication and review of data (NHMRC, 1999). In the current context, 'treatment' includes consumption of food or the property of food.
Bioavailability	The proportion of a food component such as a nutrient that is readily absorbed from the gastrointestinal tract, distributed and utilised in the body. Typically determined by comparing the area under the time-concentration in plasma curve with an appropriate reference.
Bioequivalence	In general, a property of food, having a technical specification and present in two or more different food matrices, can be said to be bioequivalent if their bioavailabilities, after administration of the same molar quantity, are similar to such a degree that their efficacy can be expected to be essentially the same.
Biological plausibility	Refers to a relationship that is consistent with existing biological and medical knowledge. For example, if a kinetic study revealed that the property of food was not absorbed (i.e. no systemic exposure) then it is biologically implausible for there to be any measurable biochemical effects attributable to its consumption.
Case control study	Patients with a certain outcome or disease and an appropriate group of controls without the outcome or disease are selected and then information is obtained on whether the subjects have been exposed to the factor under investigation (NHMRC, 1999).
Case series	The intervention has been used in a series of patients and the results reported, without the use of a separate control group (NHMRC, 1999).
Causality	Demonstration of causality considers the quality and quantity of original research in humans that support a beneficial effect of the food or property of the food; the strength of the association between the food and health effect (i.e. statistical significance of the effect) and the relationship between the amount of the food and the health effect (i.e. dose-response) (Health Canada, 2009).
Cochrane Collaboration	An international, non-profit, independent organisation that produces and disseminates systematic reviews of healthcare interventions, and promotes the search for evidence in the form of clinical trials and other studies of the effects of interventions (www.cochrane.org).

Cohort study	A study of groups who have been exposed, or not exposed, to the factor of interest (NHMRC, 1999). In a prospective cohort study, the groups are selected before measurement begins. In the current context, a typical prospective cohort study would measure dietary intake in all participants then divide them into different groups according to how much of the property of food was consumed (e.g. high, medium and low intake). The cohort is followed through time to determine whether there is a different incidence of the outcome among the groups.
Confounding	The measure of a treatment effect is distorted because of differences in variables between the treatment and control groups that are also related to the outcome (NHMRC, 1999). In observational studies, confounder variables must be measured and their effects removed statistically. In a trial, masked allocation and good randomisation are used to manage confounding.
Cross sectional study	A study that examines the relationship between health outcomes and other variables of interest as they exist in a defined population at one particular time (i.e. exposure and outcomes are both measured at the same time) (NHMRC, 1999). National nutrition surveys are an example.
Ecological study	A study in which those analysed are populations or groups rather than individuals, such as a study that compares disease rates in two different countries.
Exclusion criteria	Criteria used to establish the literature searches undertaken for a systematic review, by defining factors that are not to be included in retrieved studies. For example, the literature search may exclude studies where an intervention was not taken orally, or studies were only conducted in sick participants. These criteria are set before the literature search is undertaken.
General level health claim	See Standard 1.1.2 – Definitions used throughout the Code (www.foodstandards.gov.au)
Generalisability	Refers to the extent to which a study's results provide a correct basis for generalisation beyond the setting of the study and the particular people studied. It implies the application of the results of a study to another group or population (NHMRC, 1999)
GRADE	GRADE (Grading of Recommendations Assessment, Development and Evaluation) provides a framework for assessing quality that encourages transparency and an explicit accounting of the judgments made. GRADE distinguishes between quality assessment conducted as part of a systematic review and that undertaken as part of guideline development (Balshem et al, 2011).

Health effect	See Standard 1.1.2 – Definitions used throughout the Code (www.foodstandards.gov.au)
High level health claim	See Standard 1.1.2 – Definitions used throughout the Code (www.foodstandards.gov.au)
Inclusion criteria	Criteria used to establish the literature searches undertaken for a systematic review, by defining factors such as the study design, intervention and population groups that must be present in the studies selected. These criteria are set before the literature search is undertaken. Also sometimes referred to as eligibility criteria.
<i>In vitro</i>	Studies conducted in isolated biological material, rather than in whole, living organisms. <i>In vitro</i> studies are not suitable for establishing food health relationships.
Meta-analysis	The combination of results from studies identified in a literature review to derive an overall result. Preferably, the studies should be identified from a systematic search of the literature, not a haphazard search. Meta-regression is a type of meta-analysis that examines dose-response.
Observational study	Studies in which the researchers observe and measure what people are doing or what happens to them. It is a general term that includes cohort studies, case-control studies, cross-sectional studies, case-series and ecological studies. Sometimes also known as epidemiological studies. In contrast, an experimental study is one in which the researchers change what is happening to people; a randomised controlled trial is the best type of experimental study.
Original literature	Reports of individual observational (e.g. cohort) and experimental (e.g. randomised controlled trials) studies. Does not include reviews of a group of studies.
Systematic review	A high-level overview of primary research (i.e. original literature) on a particular research question that tries to identify, select, synthesize and appraise all high quality research evidence relevant to that question in order to answer it (www.cochrane.org).

Appendix 1: Template for a PRISMA flow diagram documenting filtering of studies retrieved in a literature search

(from Moher et al, 2009 doi:10.1371/journal.pmed.1000097.g001)



Appendix 2: Examples of table layout for showing study characteristics

An example of a summary table of study characteristics

Study reference	Study design	Study objectives	Sample size & loss to follow up	Characteristics of participants	Method to measure food consumption	Confounders measured	Method used to measure health effect	Study results ¹ (including effect size and statistical significance)	Adverse effects noted
Zones et al 2000	RCT, unblinded, primary prevention experimental trial. Comments: Further detail required on randomisation technique	Increase consumption of fruits and vegetables to 7 serves/day will result in beneficial changes in plasma lipid concentration.	N=85	US white males 19-69 years, 62 US white females aged 18-63 years. Inclusion: Eat ≤ 3 serves fruit & vegetable per day Total cholesterol < 6mmol/L Exclusion: Use of lipid lowering medication BMI > 30 Duration: 8 week test, plus 2 week run-in Comments: Statistical power adequate, but longer duration would have assisted study weight.	Diet measured with 2x4day diet records (wk 0 and 4) & 1x24 hr recall (wk6). Comments: Needed additional diet measurement at end of study. Did not fully define what was included as fruit and vegetable, e.g. inclusion of processed varieties. Bioavailability not relevant.	Changes in antioxidant intake. Decreased total and saturated fat intake and increased carbohydrate intake. Body mass increase within energy intake not controlled. Comments: no adjustment for confounders undertaken.	Plasma lipids (HDL-, LDL-cholesterol) Comments: Outcome measured with appropriate methods in experienced lab with established QC procedures.	No relationship identified between consumption of an extra 4 serves of fruit and vegetables per day for 8 weeks, and serum LDL- and HDL-cholesterol levels in healthy, non-obese adult males and females. No statistical significant relationships found ($p > 0.05$) for all outcomes).	Weight gain in some participant when food and vegetable were added to existing food consumption. (mean gain – 1kg)

An example of a summary table for study results.

Study results 2					
Intake	Baseline control	Baseline test	Wk 4 control	Wk 4 test	Adjusted difference* (95%CI)
Fruit (g)	37±51	93±118	55±84	256±132	177 (124-225)
Vegetables (g)	196±87	228±127	218±104	332±149	104(45-160)
Fibre (g)	17	19	19	25	6.2(2.1-9.0)

*Between treatment and control groups at week 4 adjusted for age, sex, baseline value.

Plasma lipid concentration (mmol/L) (mean±SD)

Lipid	B'line control	B'line test	Wk 8 control	Wk 8 test	Adj. diff (95% CI)
LDL	3.17±0.85	2.95±0.91	2.97±0.92	2.82±0.85	0.02(-0.29-0.25)
HDL	1.27±0.38	1.18±0.38	1.35±0.40	1.23±0.41	-0.08(-0.15-0.001)

Appendix 3: Examples of quality appraisal tools

Appendix 3 shows the quality appraisal tool suggested by Health Canada for experimental studies and observational studies. This is one approach to achieve some of the requirements of paragraph S6—2(e) of Schedule 6.

Table 3.1: An example of a quality appraisal tool for experimental studies (Health Canada 2009, reproduced with permission)

Reference (Author, year):

Assign a score of 1 for each 'Yes', and a score of 0 for each 'No/NR'.

Item	Question	Score	
		YES (1)	NO /NR(0)
1. Inclusion/exclusion criteria	Were the inclusion and exclusion criteria for study participation reported? (eg. Age greater than 50 years, no history of heart disease)?		
2. Group allocation ¹	Was the study described as randomized?		
	Was the randomization method reported?		
	Was the randomization appropriate? ²		
3. Blinding	Was the allocation concealed? ³		
	Were the study subjects blinded to the intervention received?		
4. Attrition	Were the researcher personnel blinded to the intervention received by the subjects?		
	Were attrition numerically reported?		
5. Exposure/intervention	Were the reasons for withdrawals and dropouts provided? ⁴		
	Was the type of food described (eg. Composition, matrix)?		
6. Health effect	Was the amount of food described (i.e. dose)?		
	Was the methodology used to measure the health effect reported?		
7. Statistical analysis	Was between group statistical analysis of the health effect reported?		
	Was an intention-to-treat analysis conducted? ⁵		
8. Potential confounders	Were potential confounders of the food health relationship considered? ⁶		
TOTAL SCORE (maximum of 15)			
Higher quality	(score 8-15)		
Lower quality	(score 0-7)		

*Notes: NR=Not reported

¹Studies without an appropriate control group would be excluded at Step of applying inclusion and exclusion criteria

² Examples of appropriate randomization include the use of computer-generated random number table, while date of birth and alternate allocation are examples of inappropriate methods of randomization.

³ Allocation concealment is not the same as blinding. Allocation concealment refers to the method used to implement the random allocation sequence, e.g. numbered envelopes containing assignment. It protects the assignment sequence before and until allocation. Blinding protects the sequence after subjects have been allocated.

⁴ If the study reported no attrition (i.e. no subjects were lost to follow up, withdrew or were excluded) then reasons for withdrawal/dropouts is a "non-applicable" factor. In such circumstances, check 'YES' so as to not unfairly lose a point.

⁵ If there was no subject attrition, a per-protocol analysis is appropriate and an intention-to-treat analysis not applicable. In such a case, check 'YES' so as to not unfairly lose a point.

⁶ Confounding could have occurred during subject selection, study conduct or data analysis. If randomization is successful and between groups differences that may have occurred during study conduct are considered during statistical analysis, then confounders were considered.

Table 3.2: An example of a quality appraisal tool for observational studies, e.g. cohort and case-control studies (Health Canada 2009, reproduced with permission)

Reference (Author, year):

Assign a score of 1 for each 'Yes', and a score of 0 for each 'No/NR'.

Item	Question	Score	
		YES (1)	NO/NR (0)
1. Inclusion/exclusion criteria	Were the inclusion and exclusion criteria for study participation reported? (eg. age greater than 50 years, no history of heart disease)?		
2. Attrition	Was attrition numerically reported?		
	Were the reasons for withdrawals and dropouts provided? ¹		
3. Exposure	Was the methodology used to measure the exposure reported?		
	Was the exposure assessed more than once?		
4. Health outcome	Was the methodology used to measure the health outcome reported?		
	Was the health outcome verified (eg. Through assessment of medical records, confirmation by a health practitioner)?		
5. Blinding	Were the outcome assessors blinded to the exposure status?		
6. Baseline comparability of groups	Were the subjects in different exposure groups compared at baseline?		
7. Statistical analysis	Was the statistical significance of the trend reported? ^{2,3}		
8. Potential confounders	Were key confounders related to subjects' demographics accounted for in the statistical analysis?		
	Were key confounders related to other risk factors of the health outcome accounted for in the statistical analysis? ^{2,4}		
TOTAL SCORE (maximum of 12)			
Higher quality	(score ≥ 7)		
Lower quality	(score ≤ 6)		

*Notes: NR= not reported

¹ If the study reported no attrition (i.e. no subjects were lost to follow up, withdrew or were excluded) then reasons for withdrawal and dropout is an NA factor. In such case, check 'YES' so as to not unfairly lose a point.

² Specify the confounders considered in footer to this table. Confounding could have occurred during subject selection (e.g. inclusion/exclusion criteria), study conduct, or data analysis.

³ Confounders related to subjects' demographics include age, sex and ethnicity.

⁴ Confounders related to other risk factors of the health outcome include, but are not limited to, diet, physical activity, smoking, alcohol intake, body mass index, weight loss, health status, family history and medication/supplement use.

Appendix 4: The GRADE system

GRADE (Grading of Recommendations Assessment, Development and Evaluation) is a tool for rating the quality of evidence and the strength of recommendations (<http://www.gradeworkinggroup.org/index.htm>, accessed 28 August, 2013). In the current context, only the first part of the tool (rating the quality of evidence) is relevant. Health claims do not involve making recommendations or guidelines and so the second component of GRADE (rating the strength of a recommendation) is not relevant.

'GRADE distinguishes between quality assessment conducted as part of a systematic review and that undertaken in the process of guideline development. In the context of a systematic review, the ratings of the quality of evidence reflect the extent of our confidence that the estimates of effect are correct' (Balshem et al. 2011). In the GRADE system, the term 'quality' is used to refer to the concept of degree of certainty rather than a quality appraisal of individual studies. As noted elsewhere, terminology is not yet consistent in this field and readers need to examine the context of other authors' writings and not assume that the terminology used in Standard 1.2.7 is used by other authors.

Table 4.1 below shows both the rating approach to individual studies and then to the body of evidence of the studies as a whole. Individual studies are rated initially as high, if a randomised controlled trial, or low if observational. Depending on how well an individual study is conducted, and some other features, this initial rating can decrease (if a trial) or either increase or decrease (if observational). Finally the body of evidence on a topic is rated.

Table 4.2 shows an earlier and an updated interpretation of the quality/degree of certainty ratings for the body of evidence in the GRADE system. Details on how to use each feature listed in the tables are given in a series of articles (the GRADE Series. <http://www.jclinepi.com/content/jce-GRADE-Series> (accessed 15 March 2013)).

Table 4.1: A summary of GRADE's approach to rating quality of evidence (from Balshem et al. 2011; reproduced with permission)

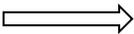
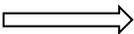
Study design	Initial quality of a body of evidence		Lower if	Higher if	Quality of a body of evidence
Randomised trials	High		Risk of Bias	Large effect	High 
			- 1 Serious - 2 Very serious	+1 Large +2 Very Large	
Observational studies	Low		Inconsistency	Dose response	Moderate 
			- 1 Serious - 2 Very serious	+ 1 Evidence of a Gradient	
			Indirectness	All plausible residual confounding	Low 
			- 1 Serious - 2 Very serious	+1 Would reduce a demonstrated effect	
Imprecision		Very low 			
			Publication bias	+1 Would suggest a spurious effect if no effect was observed	
			- 1 Serious - 2 Very serious		

Table 4.2: Interpretation of the summary designation for the quality of a body of evidence (Balshem et al. 2011, reproduced with permission)

Quality level	Current definition	Previous definition
High	We are very confident that the true effect lies close to that of the estimate of the effect	Further research is very unlikely to change our confidence in the estimate of effect
Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect	Any estimate of effect is very uncertain

Appendix 5: Revision History

Version 1.2

- this appendix was added
- disclaimer and copyright conditions revised
- some additional comments were added to Section 3.5.1 noting that available quality assessment tools do not necessarily cover all the required items, and that different quality criteria would be more or less important depending on the review characteristics
- remarks about the implications of advance online publication of papers by journals when updating existing reviews added in Section 3.8.2
- references to the Code were updated to reflect the revised Code as at 1 March 2016
- weblinks to documents were reviewed and revised as required
- minor editorial changes