

## **Continuous update literature review on diet and cancer**

### **Protocol for systematic review on nutrition, physical activity and health outcomes in breast cancer survivors – version 2.**

The Continuous Update Project on breast cancer survivors is an extension of the Continuous Update Project (CUP) on diet, nutrition, physical activity, and cancer prevention.

The current protocol for the Continuous Update on breast cancer survivors should ensure consistency of approach to the evidence used in the literature reviews for the WCRF/AICR Second Expert Report for cancer incidence and in the CUP.

The starting points for this protocol are:

- The convention for conducting systematic reviews developed by WCRF International for the Second Expert Report.<sup>1</sup>
- The recommendations of the Cancer Survivors Protocol Development Committee (Appendix 1)
- The protocol developed by the SLR group on cancer survivors for the Second Expert Report (SLR centre: University of Bristol) (Appendix 2)

The peer-reviewed protocol will represent the agreed plan for the Continuous Update on breast cancer survivors. Should departure from the agreed plan be considered necessary at a later stage, this must be agreed by the WCRF/AICR Secretariat and the reasons documented.

#### **CANCER SURVIVORS PROTOCOL DEVELOPMENT COMMITTEE MEMBERS.**

Larry Kushi (LK), Chair  
Kaiser Permanente, Oakland CA, USA

Marie-Christine Boutron-Ruault  
INSERM, E3N-EPIC Group, Institut Gustave Roussy, Villejuif (France)

Bas Bueno-de-Mesquita  
National Institute for Public Health and the Environment (RIVM), The Netherlands

Josette Chor  
School of Public Health and Primary Care, Chinese University of Hong Kong

Wendy Demark-Wahnefried  
University of Alabama at Birmingham Comprehensive Cancer Center, USA

Michelle Harvie  
University of Manchester, UK

#### **WCRF Executive**

Kate Allen  
Director, WCRF International

Deirdre McGinley-Gieser  
Senior Vice President for Programs, AICR

**Secretariat**

Martin Wiseman  
Medical and Scientific Adviser, WCRF International

Rachel Thompson,  
Science Programme Manager (Nutrition) WCRF International

Susan Higginbotham,  
Director for Research AICR

## **BACKGROUND**

The Panel of Experts for the 2007 WCRF-AICR report concluded that the available evidence from clinical trials on nutrition and physical activity and cancer prognosis was limited, and did not support specific recommendations for cancer survivors<sup>1,2</sup>. The Panel recommended that all cancer survivors should receive nutritional care from an appropriately trained professional and if able to do so, and unless otherwise advised, cancer survivors should aim to follow the recommendations for diet, healthy weight and physical activity for cancer prevention.

Advances in early detection and treatment have increased breast cancer survival considerably. With the increasing numbers of long-term survivors, research specific to cancer prognosis, new breast cancer events, quality of life and mortality is of considerable public health importance. In Europe, the five-year relative survival of women diagnosed with breast cancer in 1995–1999 is estimated to be above 82% in Northern Europe, France, Italy and Switzerland, and around 77% in the United Kingdom. In Eastern European countries, five-year relative survival is around 73% or lower<sup>3</sup>. In United States, the overall five-year relative survival for 1999-2006 has been estimated as 89% and for localized disease, the estimated five-year relative survival is 98%<sup>4</sup>.

Recent studies suggest that diet and exercise interventions may be of benefit in ameliorating adverse sequelae of cancer and its treatment, as well as cancer-specific survival and overall survival after breast cancer.<sup>5-7</sup>

The objective of this project is to identify and summarize the available information from published epidemiologic research on lifestyle and several health outcomes among women with a history of breast cancer. This review differs to the Search Literature Review for the 2007 WCRF/AICR report in two main aspects: it will be focused on studies in breast cancer survivors and it will include not only clinical trials but also follow-up studies in breast cancer survivors.

## **RESEARCH QUESTION.**

The research topic is:

The associations between food, nutrition, dietary patterns, weight control, nutrition-related complementary medicine and physical activity with mortality, breast cancer recurrence, second cancers, long-term treatment side effects and quality of life in breast cancer survivors.

### **1. REVIEW TEAM**

**REVIEW TEAM, IMPERIAL COLLEGE LONDON**

Teresa Norat, PhD  
Principal Investigator, Imperial College London

Doris Chan, MSc  
Nutritional epidemiologist, Imperial College London

Dagfinn Aune  
Nutritional epidemiologist, Imperial College London

Rui Vieira,  
Data manager, Imperial College London

Tim Reeves  
Research Support Librarian: Medicine  
Central Library, Imperial College London

**EXTERNAL COLLABORATORS**

Darren Greenwood, PhD  
Statistical advisor, University of Leeds

Rachel Thompson, PhD  
Review coordinator, WCRF

## 2. TIMELINE.

Task	Deadline
Preparation and approval of the protocol	November 2010
Changes to the structure of the database	June-December 2010
Start Medline search of relevant articles	November 2010
Review abstracts and citations identified in initial electronic search. First selection of papers for complete review	30 January 2011
Report to WCRF number of papers by study type for establishing priorities <sup>¶</sup>	28 February 2011
Select papers for data extraction <sup>†</sup>	30 May 2012
Data extraction <sup>†</sup>	30 June 2012
Data analysis	30 August 2012
Preparation of report for Panel of experts <sup>¶</sup>	30 September 2012
Send report to WCRF-AICR	30 September 2012
Transfer Endnote files to WCRF	30 September 2012

<sup>†</sup> Continue through all the review process

<sup>¶</sup> Deliveries will depend of priorities.

### 3. SEARCH STRATEGY

The search strategy will be:

- a) Search for all studies relating to breast cancer prognosis (mortality, breast cancer recurrence, second cancers, long-term treatment side effects and quality of life):

*Steps 1 to 5 were used by the SLR on breast cancer (Milan centre) to identify breast cancers. Steps 7 to 15 developed at the UK Cochrane Centre and used by the SLR on cancer survivors (Bristol University). The Imperial College CUP team expanded the search with step 16 neoplasms/rehabilitation/ to identify other dimensions.*

- b) Search for all studies relating to food, nutrition, body fatness, complementary medicine, dietary supplements and physical activity:

The CUP review team will use the search strategy developed by medical librarians and tested during the SLR for the WCRF/AICR 2<sup>nd</sup> Expert Report. The search was additionally reviewed and implemented for Ovid by a librarian at Imperial College (Appendix 3). The search strategy retrieves foods, macronutrients, micronutrients from diet and supplements, dietary supplements, herbs, breast feeding, anthropometric characteristics and physical activity.

### 4. SELECTION OF ARTICLES

Only articles that match the inclusion criteria will be updated in the database.

#### 4.1 Inclusion criteria

The articles that will be included in the systematic review:

- Investigate the associations between food, nutrition, weight control, nutrition-related complementary medicine, physical activity and mortality, breast cancer recurrence, second cancers, long-term treatment side effects and quality of life in survivors of primary breast cancer.
- The study population are pre- or post-menopausal women with diagnosis of in situ or invasive breast cancer.
- Present results of primary analysis, secondary analysis or ancillary analyses of randomized controlled trials, or follow-up studies in breast cancer survivors. If the study is a randomized trial, the trial should include at least 50 participants and the length of follow-up should be at least six months
- Report a measure of the effect/association of the intervention/exposure on the outcomes relevant to this review.

- The intervention/exposures investigated are those relevant to the WCRF/AICR 2<sup>nd</sup> Expert Report (food, nutrition, weight control, physical activity) and nutrition-related complementary medicine.
  
- Present results for any of the following outcomes:
  - Breast cancer mortality
  - Overall mortality
  - Any other mortality cause
  - Disease free survival (as defined by the authors in the identified articles)
  - Cancer recurrence
  - Second primary breast cancer
  - Other second primary cancer
  - Weight change
  
  - Quality of life (if the study is a randomized clinical trial, length of follow-up should be at least six months)
  - Development of comorbidities (e.g. fractures, cardiovascular disease, diabetes)
  - Long-term treatment related effects (e.g. lymphoedema, fatigue, osteoporosis).
  - Side-effect of diet-related modifications, physical activity interventions, nutrition-related complementary medicine, micronutrient supplementation or other dietary supplementation.
  
- Are original articles published in peer-reviewed journals.
  
- Are published in English language\*

*\* The search in this review will not be restricted by language. However, for feasibility reasons, only articles in English language will be included. Approximately 9% of clinical trials indexed in EMBASE are in languages other than English and from these about 2% are in Chinese language.*

*Articles in non-English language relevant to this review can be identified when the title and abstract are translated to English, and when the translation provides enough information to decide if the article is relevant or not to the review. The references and abstracts of relevant studies published in languages other than English will be stored in a Reference Manager database.*

*The WCRF Secretariat and the Expert Panel will decide what articles published in non-English language should be translated to English.*

## 4.2 Exclusion criteria

The articles to be excluded from the review are:

- Pooled analysis and meta-analysis (these will be used as support for interpretation, but the data will not be included in the database.)
- Comments, reviews, conference abstracts.

## 5. EXPOSURES/INTERVENTIONS.

The methods of exposure assessment will be extracted and whether the method has been validated, the number of items in the questionnaires and the number of assessments.

The duration of the exposure/intervention will be recorded as well as the time between exposure assessment/intervention and outcome assessment.

### 5.1 Labels of exposure/interventions.

During data extraction, interventions/exposures will be labelled using the exposure codes listed in the Guidelines for the search literature reviews of the 2007 WCRF/AICR expert Report<sup>1</sup>. The interventions/exposures are allocated under the main headings and subheadings listed in Appendix 4. For example, diet modifications –e.g. diets rich in fruit and vegetables and low in fats- will be coded under “Dietary patterns” and combinations of micronutrients in supplements will be coded under “Dietary Constituents”.

An additional main heading for “Nutrition-related complementary and alternative Medicine” has been added for this review (**code 9** in Appendix 4) with the following subheadings: Traditional medicine, Naturopathy, Phytotherapy, Homeopathy. Biomarkers of exposure will be extracted under the heading of the corresponding exposure, Biomarkers for which there is no evidence on appropriate validity and repeatability will not be included in the review (List of biomarkers is in Appendix 5).

### 5.2 Timeframe of exposure assessment.

The timeframe of exposure assessment in observational studies will be recorded as follows:

-Exposure assessment refers to a period **before** primary breast cancer diagnosis (childhood, adolescence, adulthood).

-Exposure assessment refers to the period **during therapy** for primary breast cancer.

-Exposure assessment refers to a period **after** primary breast cancer diagnosis.

## 6. OUTCOME



The outcomes relevant to this review are:

- Mortality
  - All cause mortality
  - Breast cancer mortality
  - Other causes of deaths
- Disease free survival (as defined by the authors in the identified articles)
- Cancer recurrence
- Second primary breast cancer
- Other second primary cancer
- Weight change
- Quality of life [psychological well being (e.g. fatigue, depression) and function (including performance status) but not spirituality].
- Treatment side effects such as lymphoedema, fatigue.
- Development of comorbidities. This includes bone health (e.g. fractures, cardiovascular disease, diabetes).
- Side-effect of diet-related modifications, physical activity interventions, nutrition-related complementary medicine, micronutrient supplementation or other dietary supplementation.

There will not be specific search for markers of tumor biology (e.g. proliferation rate, apoptosis, circulating cancer cells) because they are not relevant outcomes of the review. Results on markers of tumor biology will be extracted under “Notes” only from articles that provide results on the relevant outcomes.

## **7. DATABASES**

The databases to be searched are:

a) Medline.

b) The Cochrane Library:

- CDSR (Cochrane Database of Systematic Reviews): includes all Cochrane Reviews (and protocols) prepared by Cochrane Review Groups in The Cochrane Collaboration.

- CENTRAL (The Cochrane Central Register of Controlled Trials): is comprised of a merge of relevant records retrieved from MEDLINE, relevant records retrieved from EMBASE, all Review Groups' Specialised Registers and the hand search results register.

c) EMBASE

## **8. HAND SEARCHING FOR CITED REFERENCES**

For feasibility reasons, journals will not be hand searched.

The CUP team will review the references of meta-analyses, reviews and pooling projects identified during the search.

## **9. REFERENCE MANAGER FILES**

Reference Manager files are generated in the continuous update containing the references of the initial searches in all databases.

- 1) One of the customized fields (User Def 1) is named 'inclusion' and this field is marked 'included', 'excluded' for each paper, thereby indicating which papers are deemed potentially relevant based on an assessment of the title and abstract.
- 2) One of the customized fields (User Def 2) is named 'reasons' and this field should include the reason for exclusion for each paper.
- 3) The study identifier should be entered under the field titled 'label'.
- 4) One of the customized fields (User Def 3) is named "study design". This field indicates the study design of each paper:
  - Randomized controlled trials excluding interventions during cancer treatment.
  - Randomized controlled trials during cancer treatment.
  - Group Intervention trials
  - Observational studies where exposure refers to the period before breast cancer diagnosis
  - Observational studies where exposure refers to the period from diagnosis through adjuvant treatment.
  - Observational studies where exposure refers to the period after breast cancer diagnosis after adjuvant treatment.

The Reference Management databases will be converted to EndNote and sent to WCRF Secretariat as part of the report.

## **10. RETRIEVING ARTICLES**

The references of articles retrieved in the searches in the different databases will be merged by the database manager into a Reference Manager (RefMan) database.

Animal and in vitro studies will be excluded with the following stop terms: transgenic, mice, hamster, rat, dog, cat, in vitro. *(This procedure was tested by the SLR team Leeds during the SLR for the 2007 WCRF/AICR expert report.)*

Non-relevant exposures under the Mesh term “Complementary medicine” will be excluded using the following stop terms: Acupuncture Therapy, Anthroposophy , Auriculotherapy, Holistic Health, Mind-Body Therapies, Musculoskeletal Manipulations, Organotherapy, Reflexotherapy, Rejuvenation, Sensory Art Therapies, Speleotherapy, Spiritual Therapies, Shamanism, Aromatherapy , Eclecticism, Historical.

The database manager will identify and eliminate duplicates in the RefMan database using as key terms the first author name, publication year, journal name, volume, starting page number of the article. Automatic searches for duplicates in Ref Man are not recommended because the references retrieved in each database may be exported differently.

The reviewer will assess relevant articles on the Reference Manager database upon reading of titles and abstracts. The complete papers of relevant and potentially relevant references and of references that cannot be excluded upon reading the title and abstracts will be reviewed. A second assessment will be done after review of the complete papers.

The assessments of inclusion of articles will be done in duplicate by two independent reviewers for articles published in 2009 and 2010. If there is full agreement in the selection, 10% of the remaining articles will be double assessed for inclusion. This decision is based on feasibility of the project with the existing resources. The WCRF secretariat and the Expert Panel will be consulted on this before changes to the protocol are implemented.

## **11. LABELLING OF ARTICLES**

For consistency with the previous data collected during the SLR process for the Second Expert Report, the CUP review team will use the same labelling of articles: the unique identifier for a particular reference will be constructed using S to indicate “survivors” and a 2-letter code to represent the cancer site (e.g. BR for breast cancer), followed by a 5-digit number that will be allocated in sequence.

## **12. DATA EXTRACTION**

Data extraction will be performed by the reviewer using a screen extraction form designed by the database manager of the CUP. Extractions will be double checked by a second reviewer for 10% of the extracted articles by the first reviewer.

The data will be extracted to the WCRF database located in a protected server at Imperial College London. The structure of the existing database will be adapted to the scope of the search on breast cancer survivors before the start of the search. Further modifications of the database structure may be needed during the search.

### **12. 1. Information to be extracted.**

The list of study variables for observational and intervention studies in the CUP database is in Appendix 6.

For this review, new variables will be added:

-Study type:

- Intervention study
- Follow-up study on breast cancer survivors

-Characteristics of primary breast cancer:

- Distribution of “in situ” and invasive breast cancer in the study population
- Proportion of cases in which primary breast cancer was detected by screening.
- Distribution of the study population by stage at diagnosis
- Distribution by estrogen receptor (ER), progesterone receptor (PR), human epithelial growth factor receptor 2 (Her2) status
- Distribution by cancer subtype defined by immunohistochemical analysis or gene expression profiling (e.g. Luminal A, Luminal B, etc as given in the manuscript)
- Distribution of the study population by treatment for primary breast cancer (surgery, chemotherapy, radiotherapy, hormonal therapy specifying if tamoxifen, aromatase inhibitors, monoclonal antibodies such as herceptin, other treatments, unknown).

### **12.2 Choice of Result**

The results for all relevant exposures and outcomes will be extracted. In epidemiologic studies, authors often present a series of models, e.g. unadjusted, age-adjusted, multivariable adjusted models. Sometimes authors do additional adjustments for factors likely to be in the causal pathway (“mechanistic models”). The extracted results will be labeled depending on the model as: not adjusted, intermediately adjusted, “fully” adjusted, or mechanistic model. “Fully” adjusted models and “mechanistic” models will be extracted in this review. A “fully” adjusted model will be considered the most adjusted model in the paper that is not a “mechanistic model”. If only an unadjusted or an age-adjusted model is given in the paper, this should be extracted.

The reviewer should indicate a “best model” for inclusion in reports and meta-analyses. Usually, the “fully” adjusted model will be considered the “best model”. In there is a “mechanistic model”, the “best model” for analysis will be the “fully” adjusted and not the “mechanistic” model.

The results for subgroup and stratified analyses will be extracted and the models labelled as indicated before. The “best model” for analysis will be indicated by the reviewer.

The authors of the papers will not be contacted during the process of data extraction. Only the data provided in the article will be extracted to the database.

### **12.3 Multiple articles**

Data should be extracted for each individual article, even if there is more than one article from any one study, unless the information is identical. The most appropriate set of data on a particular exposure will be selected among the articles published on a study to ensure there is no duplication of data from the same study in an analysis.

### **12.4 Quality control**

Inclusion assessment will be done in duplicate for articles published in 2009 and 2010. If there is concordance in the selection between the two reviewers, the quality control of the selection procedure will be done by a second reviewer on only 10% of the papers excluded by the first reviewer. This is due to limited resources (section 9 “Retrieving Articles”). Any disagreement between reviewers will be solved with the principal investigator at Imperial College. In case of doubt about the study selection, the WCRF Secretariat will be contacted for advice. When discrepancies are detected, the protocol will be revised to add more clarifications.

Data extraction will be checked by a second reviewer. Only 10% of the data extraction will be reviewed. If there are discrepancies, another 10% of the extracted information will be checked.

### **12.5 Gene-nutrient interaction**

No attempt was made to critically appraise or analyse the studies that reported gene-nutrient interactions in the 2007 WCRF/AICR second expert report<sup>1</sup>. The results of relevant studies on gene-environment interactions will be described in a narrative review

## **13. ASSESSMENT OF STUDY QUALITY AND SUSCEPTIBILITY TO BIAS.**

The evaluation of randomized controlled trials will be based in the checklist proposed by the Cochrane Collaboration (<http://www.cochrane-handbook.org/>).

The dimensions of quality and susceptibility to bias in the check lists are:

- Selection bias: Systematic differences between baseline characteristics of the groups that are compared.
- Performance bias: Systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest.
- Attrition bias: Systematic differences between groups in withdrawals from a study.
- Detection bias: Systematic differences between groups in how outcomes are determined.
- Reporting bias: Systematic differences between reported and unreported findings.

The items will receive score 1 point if susceptibility to bias is low and 0 if susceptibility to bias is considered high. The total score of the article will be the sum of the item scores (details in Appendix 7).

Numerous tools have been proposed for evaluation of methodological quality of observational epidemiological studies but there is no agreed “gold standard”<sup>8</sup>. We will assess the quality of observational studies using the Newcastle-Ottawa quality assessment scale, which is simpler to use and has been used in recently published meta-analysis

([http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp)).

The dimensions included are:

- Selection of study population (ascertainment of exposure is included in this dimension).
- Comparability: control for confounding
- Outcome: ascertainment and follow-up

We will exclude the item “representativeness of the study population” as criteria of study quality, because it does not affect the study internal validity<sup>8</sup>. The characteristic of the study population will be extracted and could be use for further analysis.

Studies will not be excluded on the basis of study quality. The assessment of study quality will be used to inform narrative reviews and for sensitivity and meta-regression analyses. Scores of study quality or susceptibility to bias will be included in tables of study characteristics in the reports.

## **15. DATA ANALYSIS**

Meta-analyses and narrative reviews will complement each other.

### **15.1 When to do a meta-analysis**

A meta-analysis for a particular exposure/intervention and outcome will be conducted when three or more trials or observational studies that can be combined have been published. If meta-analyses are not possible, the results will be summarized in a narrative review.

Special care will be taken to avoid including more than once the results of the same study (e.g. previous analyses and re-analyses after a longer follow-up).

## **15. 2 Methods**

The methods that will be used to do meta-analyses will be the same methods used for the Second Expert Report<sup>1</sup>.

Meta-analyses will be conducted separately by study type, outcome and timeframe of exposure (before diagnosis, during treatment, after treatment). The best model (most often the “fully” adjusted measure of association or effect) from each analysis will be used.

In trials with multiple intervention arms and intervention of different types (e.g. one multivitamin supplement and one dietary counselling intervention), each arm will be compared with the usual treatment group (or specific placebo group) and analyzed separately. Consequently, some studies may contribute data to more than one analysis.

When multiple interventions in a trial are of the same type, the results of each arm will be compared first with the results of the control data arm separately. If the results of each arm are consistent in size and direction of effect, the data from the interventions arms will be treated as one group. This method will avoid the control groups being included twice in the same meta-analysis<sup>2</sup>. Factorial trials will be analyzed by assuming no interaction between interventions.

In meta-analysis of two categories (or “high-low” comparisons), summary RR estimates with their corresponding 95% CIs will be derived using fixed and random effect models<sup>9</sup>. A difference in the point estimate in fixed and random effect analysis must indicate that results from smaller studies differ from those of larger studies.

To estimate the dose-response relationship, category-specific risk estimates will be transformed into estimates of the relative risk (RR) associated with a unit of increase in exposure by use of the method of generalised least-squares for trend estimation<sup>10</sup>. When exposure levels are reported as means or medians for each category of exposure, these values will be used directly in the dose-response meta-analyses. If the exposure is given as an interval, the mid-point of the interval will be assigned to each closed-ended category of exposure. The median will be assigned to each open-ended category. The

median will be calculated assuming a normal distribution for exposure<sup>11</sup>. When categorical and continuous results are provided, the continuous results will be used in the dose-response meta-analysis. The relative risk estimates for each unit of increase of the exposure from each study will be combined by use of fixed and random-effect meta-analysis<sup>9</sup>.

Forest plots will be examined as usual method of assessing and displaying heterogeneity between studies. Heterogeneity will be tested using the Q statistic. The amount of heterogeneity in each meta-analysis will be quantified with the  $I^2$  statistic<sup>12</sup>. Influence-analyses to assess the effect of each study on the summary size effect estimates<sup>13</sup>. Publication and small study bias will be examined in funnel plots.

If the number of studies allows it, the sources of heterogeneity will be explored with the use of meta-regression. Possible variables to be examined are breast cancer subtype, geographic area where the study was conducted, publication year, stage of disease, duration of follow-up, timeframe of exposure assessment. Other variables that may be considered as source of heterogeneity are characterisation of the exposure (FFQ, recall, diary, self-reported or measured anthropometry etc.) and adjustment for confounders. In clinical trials, variables to be considered are whether the outcome was the primary or secondary outcome or an ancillary analysis. The interpretation of these analyses should be cautious. If a considerable number of study characteristics are considered as possible explanations for heterogeneity in a meta-analysis containing only a small number of studies, then there is a high probability that one or more will be found to explain heterogeneity, even in the absence of real associations.

The analysis will be done using STATA version 9.2 (College Station, TX, USA).

### **15.3 Missing values**

Failure to include all available evidence in the meta-analyses will reduce precision of summary estimates and may also lead to bias if propensity to report results in sufficient detail is associated with the magnitude and/or direction of associations. Published standard procedures<sup>14</sup> will be used to calculate missing information (Appendix 8).

## **16. REPORTS**

Content of the report:

- 1.** Changes to the agreed protocol
- 2.** Narrative summary of the results of the search and the data analysis



### 3. Results of the search.

- Flow chart showing number of records downloaded, number of papers thought potentially relevant after reading titles and abstracts and number of included relevant papers. The reasons for excluding papers should also be described.
- For each intervention, number of trials by outcome.
- For each exposure, number of studies by study type and outcome.

### 4. Tabulation of study characteristics

Information on the characteristics (e.g. population, exposure/intervention, outcome, study design) and results of the study (e.g. direction and magnitude) of the new studies should be summarised in tables using the same format as for the SLR for the Second Expert Report<sup>1</sup>. The tables will include the scores of study quality.

The tables for randomization controlled trials will be ordered by exposure as follows:

- Food-related interventions
- Micronutrient supplementation
- Physical activity-related interventions
- Nutrition-related complementary medicine.
- Combination of interventions

The tables of study characteristic of clinical trials will include the following information:

- Trial reference, year
- Characteristics of study population (age, race/ethnicity, BMI, menopausal status, BRCA1-2 carrier)
- Characteristics of the tumour (stage, subtype, hormone receptor status)
- Treatment at time of intervention (after, during, unclear)
- Randomization, blinding
- Intervention, duration
- Follow-up time
- Number of events and total number of participants in intervention and control arms
- Percentage of missing outcome data
- Outcome
- Results and whether these are primary endpoints or secondary endpoints; final or interim analysis; ad-hoc analysis; based on intention-to treat analysis or treated;
- Matching criteria, adjustment factors in the analysis
- Quality score

The tables for observational studies will be ordered by exposure and exposure assessment timeframe (before breast cancer diagnosis, during treatment or after treatment). The tables will contain the following information:

- Study reference, year
- Study design
- Characteristics of study population (age, ethnicity, BMI, menopausal status, use of HRT before cancer diagnosis)
- Number of cases and study size
- Whether exposure from foods or supplements, levels or increment
- Outcome
- Results
- Adjustment factors in the analysis
- Quality score

#### **5. Description of results of assessment of quality and risk of bias of included studies**

Tabulation of results for individual items of the check lists.

#### **6. Results of meta-analysis**

The results of meta-analysis will be displayed in tables and forest plots. The characteristic of excluded studies and reasons for exclusions will be tabulated.

Funnel plots for examining publication and small study bias will be included.

#### **6. Reference list.**

List of all relevant studies identified in the review.

## References

1. World Cancer Research Fund/ American Institute for Cancer Research. Systematic Literature Review – Support Resource. In : Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective. Washington DC: AICR , 2007
2. Davies AA, Davey Smith G, Harbord R, Bekkering GE, Sterne JA, Beynon R, Thomas S. Nutritional interventions and outcome in patients with cancer or preinvasive lesions: systematic review. *J Natl Cancer Inst.* 2006 Jul 19;98(14):961-73.
3. Sant M, Allemani C, Santaquilani M, Knijn A, Marchesi F, Capocaccia R; EURO CARE Working Group. EURO CARE-4. Survival of cancer patients diagnosed in 1995-1999. Results and commentary. *Eur J Cancer.* 2009 Apr;45(6):931-91 .
4. American Cancer Society. *Breast Cancer Facts & Figures 2009-2010*. Atlanta: American Cancer Society, Inc. ([http://www.cancer.org/downloads/STT/F861009\\_final%209-08-09.pdf](http://www.cancer.org/downloads/STT/F861009_final%209-08-09.pdf))
5. Demark-Wahnefried W, Morey MC, Sloane R, Snyder DC, Cohen HJ. Promoting healthy lifestyles in older cancer survivors to improve health and preserve function. *J Am Geriatr Soc.* 2009 Nov;57 Suppl 2:S262-4.
6. Kushi LH, Kwan ML, Lee MM, Ambrosone CB. Lifestyle factors and survival in women with breast cancer. *J Nutr.* 2007 Jan;137(1 Suppl):236S-242S
7. Patterson RE, Cadmus LA, Emond JA, Pierce JP. Physical activity, diet, adiposity and female breast cancer prognosis: a review of the epidemiologic literature. *Maturitas.* 2010;66(1):5-15
8. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol.* 2007 Jun;36(3):666-76.
9. DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Control Clin Trials* 7:177–188
10. N Orsini, R Bellocco and S Greenland, Generalized least squares for trend estimation of summarized dose-response data, *Stata J* 6 (2006), pp. 40–57
11. Chêne G, Thompson SG. Methods for summarizing the risk associations of quantitative variables in epidemiologic studies in a consistent form. *Am J Epidemiol.* 1996;144(6):610-21.
12. JP Higgins and SG Thompson, Quantifying heterogeneity in a meta-analysis, *Stat Med* 21 (2002), pp. 1539–1558.
13. A Tobias, Assessing the influence of a single study in meta-analysis, *Stata Tech Bull* 47 (1999), pp. 15–17.
14. Bekkering GE et al. How much of the data published in observational studies of the association between diet and colorectal or bladder cancer is usable for meta-analysis? *Am J Epidemiol* (2008);167(9):1017-26.

## **Appendix 1: The recommendations of the Cancer Survivors Protocol Development Committee**

### **Population characteristics and subject eligibility to consider**

#### **Inclusion criteria**

- In situ cancer
- Record histological type; DCIS/LCIS/other;
- Report separately from invasive cancer if possible
- Historical studies
- Historical studies with data before treatment were considered valuable. DG said that an assessment of heterogeneity would be useful to determine whether historical studies reported different results from more recent studies.
- Stage of diagnosis – Include all studies with data from diagnosis onwards

#### **Exclusion criteria**

- Male breast cancer eg mammographic density
- Pre-cancerous conditions – Exclude for breast cancer
  - May be important for other cancers
- Studies that report on a combination of different cancers ie. Breast and other cancers, without separate analyses on breast cancer survivors
  - Record these papers as they may be used in the future but do not include in this analysis

#### **Information to extract if reported in papers**

- Stage at diagnosis
  - Include information on stage and grade of cancer
- Hormone receptor status
  - Extract information on hormone receptor status
- Treatment status
  - To include type of treatment. Need to develop broad groupings e.g. hormone therapy, radiotherapy, chemotherapy (adjuvant and neoadjuvant) and surgery, and combinations.
  - Duration of treatment and timing in relation to exposure assessment
- How cancer was detected (eg, screening)
  - People attending screening may be different and how cancer is detected affects prognosis so these details should be extracted.
- BRCA1, BRCA2 and molecular subtypes
- Menopausal status
- Age
- Country study carried out in
  - BMI range was considered important as heavier women might have different effects; therefore studies in Asia might get effect at lower BMI
- Pre-diagnosis factors

- Include HRT use, diabetes status and treatment

### **Study designs to include (exclude everything else)**

- Follow up of cases from case-control studies
- Follow up of cases from cohort studies
- Cancer survivor cohorts (explicitly designed)
- Randomized controlled trials (RCTs)

### **Exposures**

#### **To include**

- List of exposures used in previous SLRs (including body size measures and change in BMI)
- Supplements of substances found in diet e.g. vitamins (to include high dose vitamins) and soy
- Biomarkers of dietary intake
- Physical activity
- Metabolic biomarkers
  - These biomarkers might be important for interpretation of the data so extract these if reported in included papers eg IGF

#### **To exclude**

- Herbal products
- Complementary and alternative medicine (CAM)
  - There are a lot of agents tested in small populations. It was decided to restrict the products to those than are found in ordinary diets.
  - These products need to be taken into account in analyses so information on use of such products should be extracted if reported in papers.
- Nutrition support (e.g. enteral/parenteral nutrition)
  - Exclude as an exposure but extract if papers report on nutrition support as it may be a modifier

#### **Timing of exposure to include**

- Pre-diagnosis
- Post-diagnosis
- Multiple time points
  - Extract information on timing of exposure in relation to time before/after diagnosis

#### **Endpoints/outcomes to include**

- All-cause or cancer (all types) mortality
- Breast-cancer specific mortality
- Disease-free survival
  - It was questioned how we would define this. It was decided that anything survival related that could be understood in papers should be extracted.

- Recurrence-free survival
- Recurrence
- New primary cancer (breast or other)
  - Breast cancer survivors are at increased risk of second primary breast and other cancers. This is important information; however will be unlikely to be able to analyse by specific cancers.
- Quality of life outcomes
  - Include only intervention studies with attention control/comparison groups that follow-up participants for at least 6 months. Short-term studies were seen as difficult to interpret.
  - It was decided not to restrict studies based on sample size
  - Includes fatigue, depression, functional/performance status
  - There was some concern that searches may pick up huge numbers of potential and/or relevant papers and that the criteria for inclusion for quality of life might need to be modified once the searches are under way.
  - A specialist in searching quality of life data should be consulted.
- Comorbidity
  - Information on other diseases (e.g. cardiovascular disease, osteoporosis, fractures) should be extracted.
- Lymphoedema

### **Study quality**

- classify studies by extent to which measures of key confounders and covariates are measured
  - Extract information on randomization method, blinding, concealment and intention to treat analysis for randomized controlled trials.

# **Systematic Review Protocol**

## **(Diet, Exercise and Cancer Survival)**

### **1. Research question**

Dietary modifications and exercise interventions in the management of cancer.

### **2. Reviewers**

**George Davey Smith, Professor of Clinical Epidemiology**

University of Bristol, Department of Social Medicine, Canynge Hall, Whiteladies Road, Bristol, UK.

**Anna Davies, Research Associate in Epidemiology**

University of Bristol, Department of Social Medicine, Canynge Hall, Whiteladies Road, Bristol, UK.

**Roger Harbord, Research Associate in Medical Statistics**

University of Bristol, Department of Social Medicine, Canynge Hall, Whiteladies Road, Bristol, UK.

**Jonathan Sterne, Senior Lecturer in Medical Statistics**

University of Bristol, Department of Social Medicine, Canynge Hall, Whiteladies Road, Bristol, UK.

**Steve Thomas, Consultant Surgeon and Senior Lecturer**

Department of Oral and Dental Science, Division of Oral and Maxillofacial Surgery, University of Bristol, Bristol, UK.

### **3. Sources of support**

World Cancer Research Fund

### **4. Objectives**

To investigate whether dietary modification and exercise have a role in the management of, and ultimately survival from, cancer.

### **5. Background**

In 1996 there were over 10 million new cases of cancer worldwide and just over 7 million cancer deaths (WHO, 1997). Traditionally, cancer has been labelled as a disease of developed countries. Data suggests, however, there is also an increasing prevalence of the disease in economically developing countries, in particular in urban areas. These figures suggest that cancer is a matter for public health concern worldwide (WHO, 1997).

The role of diet and exercise in the prevention of cancer has been extensively studied. The literature in this field was reviewed by a recent WCRF / AICR report: *Food, nutrition and the Prevention of Cancer: a global perspective* (WCRF and AICR, 1997). By comparison, there is considerably less scientific research on the role diet and exercise may play in the treatment of cancer. This means that there is greater uncertainty, for both health professionals and cancer survivors, surrounding decisions to include dietary modifications and / or exercise interventions as part of a holistic approach to cancer management (AICR, 2002a; AICR, 2002b).

The role of biologically active components of the diet at various stages of the carcinogenic process is summarized in the WCRF/AICR report (WCRF and AICR, 1997). This, together with the reported evidence from epidemiological studies, has led to a series of recommendations for changes in diet and activity levels to reduce primary cancer risk (AICR, 2002a). For cancer survivors the guidelines are less clear: the AICR state that ‘...most experts seem to agree that cancer survivors should consider research results regarding risk reduction for primary cancers as being relevant to their situation’ (AICR, 2002a). A systematic review of the literature is needed to further clarify this

statement and the association between diet, exercise and cancer survival. To address this we intend to carry out two reviews:

**Review 1: Systematic review of randomized controlled trials (RCTs)** investigating the effect of dietary modification and exercise intervention on cancer survival

**Review 2: Finding and updating the most recent epidemiological reviews** on the association between diet, exercise and cancer survival for each cancer site.

**Review 1:** Systematic review of randomised control trials investigating the effect of dietary modification and exercise intervention on cancer survival.

## **1. Criteria For Considering Studies For This Review**

### **Type of studies**

All randomized controlled trials (RCTs).

### **Type of participants**

Males and females of any age with a diagnosis of cancer.

### **Study selection criteria**

- Dietary modifications as a result of dietary life-style changes, dietary education, micronutrient supplementation and complementary medicine.
- Diet and life-style modifications consequent on the disease or its treatment will NOT be included. These include peri- and post-operative dietary modifications together with calorie enhancement for cancer cachexia.
- All exercise interventions will be included.
- Dietary modifications and exercise interventions will be included regardless of their duration or the route of dietary intake used. All concomitant interventions will be included.

### **Types of outcomes**

- survival / all-cause mortality
- cancer mortality
- primary cancer recurrence
- second primary cancer
- quality of life
- side effects

## **2. Search strategy**

All search strategies will be generated with the consultation of a medical librarian. The searches will not be limited to English. We will attempt to get translations of all non-English papers, which appear relevant. A second party will not independently verify translations. As far as possible, translators will be 'blinded' to the exact nature of the study. Only literature published in peer review journals will be included in the review (i.e. no 'grey literature' will be searched). Searching will be carried out using the following sources and time-periods:

- The Cochrane Library (2003, Issue 2). Searches will include: DARE (Database of Abstracts of Reviews of Effects); CDSR (Cochrane Database of Systematic Reviews); HTA (Health Technology Assessment) and CENTRAL (The Cochrane Central Register of Controlled Trials).
- MEDLINE (2000-present). Medline has already been systematically searched - using a highly sensitive search strategy developed by Carol Lefebvre and colleagues at the UK Cochrane Centre (Dickersin K et al, 1994) – to identify all RCTs up to and including the year



1999. The identified RCTs are listed in the Cochrane Central Register of Controlled Trials (CENTRAL) (found in the Cochrane Library, listed above as a source to search). For this reason we intend to only search Medline from 2000-2003.

- EMBASE (1980 – present)
- ISI Web of Science (1981 – present)
- BIOSIS (Previews) (1985 – present)
- AMED (Allied and Complementary Medicine Database) (1981 – present)
- Follow-up of references from relevant papers
- Personal communication with experts

### **Search strategy for MEDLINE**

This search strategy will be adapted for use in all other databases.

#### **a) Searching for all studies relating to cancer and survival**

- 1 exp neoplasms/
- 2 cancer\$.tw.
- 3 neoplasm\$.tw.
- 4 or/1-3**
- 5 survivors/
- 6 exp survival analysis/
- 7 surviv\$.tw.
- 8 recurrence/
- 9 recur\$.tw.
- 10 quality of life/ or quality of life.tw. or qaly\$.tw.
- 11 mortality/
- 12 survival rate/
- 13 (manag\$ adj3 cancer\$.tw
- 14 or/5-13**
- 15 4 and 14**

#### **b) Searching for all studies relating to dietary modification**

- 16 exp "food and beverages"/
- 17 food\$.tw.
- 18 supplement\$.tw.
- 19 exp diet/
- 20 exp diet therapy/
- 21 diet\$.tw.
- 22 exp nutrition/
- 23 nutri\$.tw.
- 24 exp dietary fats/
- 25 exp dietary proteins/
- 26 exp dietary carbohydrates/
- 27 exp vitamins/
- 28 exp Feeding Behavior/
- 29 exp Drinking Behavior/
- 30 exp trace elements/
- 31 exp antioxidants/
- 32 exp Micronutrients/
- 33 calcium, dietary/
- 34 phosphorus, dietary/
- 35 exp sodium, dietary/

- 36 potassium, dietary/
- 37 iron, dietary/
- 38 or/16-37**

**c) Searching for all studies relating to exercise modification**

- 39 exp Exercise Movement Techniques/
- 40 exp exertion/
- 41 exp sports/
- 42 exp physical fitness/
- 43 exp exercise/
- 44 exercis\$.tw.
- 45 (physical\$ adj3 activ\$).tw.
- 46 (physical\$ adj3 fit\$).tw.
- 47 (physical\$ adj3 train\$).tw.
- 48 or/39-47**

**d) Filter for randomised control trials – developed by Carol Lefebvre at the UK Cochrane Centre (Dickersin K et al, 1994)**

- 49 randomized controlled trial.pt.
- 50 controlled clinical trial.pt.
- 51 RANDOMIZED CONTROLLED TRIALS.sh.
- 52 RANDOM ALLOCATION.sh.
- 53 DOUBLE BLIND METHOD.sh.
- 54 SINGLE-BLIND METHOD.sh.
- 55 or/49-54
- 56 (ANIMAL not HUMAN).sh.
- 57 55 not 56
- 58 CLINICAL TRIAL.pt.
- 59 exp CLINICAL TRIALS/
- 60 (clin\$ adj25 trial\$).ti,ab.
- 61 ((singl\$ or doubl\$ or trebl\$ or tripl\$) adj25 (blind\$ or mask\$)).ti,ab.
- 62 PLACEBOS.sh.
- 63 placebo\$.ti,ab.
- 64 random\$.ti,ab.
- 65 RESEARCH DESIGN.sh.
- 66 or/58-65
- 67 66 not 56
- 68 67 not 57
- 69 57 or 68

**e) Selecting all RCTs on cancer survival that involve dietary or exercise modification**

- 70 15 and 69 and (38 or 48)**

### **3. Methods of the review**

#### **Data Collection**

An 'In/Out' form (appendix 1) will be used to assess each study's inclusion (or otherwise) into the review. The inclusion of studies will be assessed independently by two assessors and differences between reviewer's results will be resolved by discussion and, when necessary, in consultation with a third reviewer.

A 'Data Extraction' form will be specifically designed for the review (appendix 2). Data abstraction will be performed independently by two researchers and differences between reviewer's results will be resolved by returning to the relevant literature, discussion and, when necessary, consultation with a third reviewer. Data will be collected on the following:

- Target group
- Cancer type
- Intervention type and details of how many arms
- Study Location
- Period of intervention
- Control group
- Compliance
- Experimental and control arm participant comparison (number, male:female ratio, age, ethnicity, non-compliance, number analysed for outcome measures)
- Details of any switches in treatment condition
- Outcomes (whether analysed on intention to treat basis, comparisons between experimental and control arm)
- Quality of the trial (selection bias, performance and detection bias, losses to follow up)

#### **Data Analysis**

Where appropriate differences in outcomes comparing treatment and control groups will be combined across studies using standard methods for meta-analysis. Fixed (common) effect methods will be used. Forest plots will be used to display results and to examine possible heterogeneity between studies. In addition to standard tests for heterogeneity  $I^2$  statistics will be used measure the amount of heterogeneity (Higgins et al, 2002). In the presence of heterogeneity we will compare results from fixed effect and random effects analyses.

Funnel plots will be used to explore "small study effects" (the tendency for smaller studies in a meta-analysis to show larger treatment effects) (Sterne et al, 2000). If funnel plot asymmetry is observed careful consideration will be given as to its causes as well as the possible impact, on the overall estimate of treatment effect, from any meta-analysis performed (Sterne et al, 2001a).

Where there is sufficient data, sensitivity analyses will be carried out to investigate the impact, on the summary estimate of effect, of excluding studies with inadequate or unclear allocation concealment, trials in which blinding was not adequate and /or trials in which methods for dealing with loss-to-follow-up were not adequate. (The quality assessment section of the data extraction form will assess these characteristics).

A preliminary search of the literature suggests that is highly unlikely that there will be enough studies for dose-response graphs to be plotted or for meta-regression analyses to be done. The design of RCTs means that confounding should not be an issue and so confounding will not be considered.

All analysis will be undertaken using the statistical package Stata (version 8) in which a comprehensive set of user-written commands is available for meta-analysis (Sterne et al, 2001b).

## **Review 2**

Finding and updating the most recent epidemiological reviews on the association between diet, exercise and cancer survival for each cancer site.

### **1. Criteria for considering studies for this review**

#### **Type of studies**

All case-control, cohort (retrospective and longitudinal) and cross-sectional studies.

#### **Type of participants**

Males and females of any age with a diagnosis of cancer.

#### **Study selection criteria**

- Dietary modifications as a result of dietary life-style changes, dietary education, micronutrient supplementation and complementary medicine.
- Diet and life-style modifications consequent on the disease or its treatment will NOT be included. These include peri- and post-operative dietary modifications together with calorie enhancement for cancer cachexia.
- All exercise interventions will be included.
- Dietary modifications and exercise interventions will be included regardless of their duration or the route of dietary intake used. All concomitant interventions will be included.

#### **Types of outcomes**

- survival / all-cause mortality
- cancer mortality
- primary cancer recurrence
- second primary cancer
- quality of life
- side effects

### **2. Search strategy**

All search strategies will be generated with the consultation of a medical librarian. The searches will not be limited to English. We will attempt to get translations of all non-English papers, which appear relevant. As far as possible, translators will be 'blinded' to the exact nature of the study. Only literature published in peer review journals will be included in the review (i.e. no 'grey literature' will be searched). Searching will be carried out using the following sources and time-periods:

- The Cochrane Library (2003, Issue 2). Searches will include: DARE (Database of Abstracts of Reviews of Effects); CDSR (Cochrane Database of Systematic Reviews); HTA (Health Technology Assessment)
- MEDLINE (1966-present)
- EMBASE (1980 – present)
- ISI Web of Science (1981 – present)
- BIOSIS (Previews) (1985 – present)
- AMED (Allied and Complementary Medicine Database) (1981 – present)
- Follow-up of references from relevant papers
- Personal communication with experts

The nature of this review requires two search strategies:

- 1) The first search strategy aims to identify the most recent epidemiological review (if one exists) on the association between diet and/or exercise and cancer survival for each of the cancer sites: breast, mouth, pharynx, nasopharynx, larynx, oesophagus, lung, stomach,

pancreas, gallbladder, liver, colon, rectum, ovary, endometrium, cervix, thyroid, kidney, bladder, prostate, skin, bone, nervous tissue, haematological and lymphatic cancers.

### Search Strategy for MEDLINE

This search strategy will be adapted for use in all other databases

#### a) Searching for all studies relating to cancer and survival

- 1 exp neoplasms/
- 2 cancer\$.tw.
- 3 neoplasm\$.tw.
- 4 or/1-3**
- 5 survivors/
- 6 exp survival analysis/
- 7 surviv\$.tw.
- 8 recurrence/
- 9 recur\$.tw.
- 10 quality of life/ or quality of life.tw. or qaly\$.tw.
- 11 mortality/
- 12 survival rate/
- 13 (manag\$ adj3 cancer\$).tw
- 14 or/5-13**
- 15 4 and 14**

#### b) Searching for all studies relating to dietary modification

- 16 exp "food and beverages"/
- 17 food\$.tw.
- 18 supplement\$.tw.
- 19 exp diet/
- 20 exp diet therapy/
- 21 diet\$.tw.
- 22 exp nutrition/
- 23 nutri\$.tw.
- 24 exp dietary fats/
- 25 exp dietary proteins/
- 26 exp dietary carbohydrates/
- 27 exp vitamins/
- 28 exp Feeding Behavior/
- 29 exp Drinking Behavior/
- 30 exp trace elements/
- 31 exp antioxidants/
- 32 exp Micronutrients/
- 33 calcium, dietary/
- 34 phosphorus, dietary/
- 35 exp sodium, dietary/
- 36 potassium, dietary/
- 37 iron, dietary/
- 38 or/16-37**

#### c) Searching for all studies relating to exercise modification

39 exp Exercise Movement Techniques/  
 40 exp exertion/  
 41 exp sports/  
 42 exp physical fitness/  
 43 exp exercise/  
 44 exercis\$.tw.  
 45 (physical\$ adj3 activ\$).tw.  
 46 (physical\$ adj3 fit\$).tw.  
 47 (physical\$ adj3 train\$).tw.  
**48 or/39-47**

**d) Filter for selecting only review papers: adapted from a high sensitivity search filter for systematic reviews (CRD, 2001)**

49 (review or review,tutorial or review, academic or review, literature).pt  
 50 (systematic adj review\$).tw  
 51 (data adj synthesis).tw.  
 52 (published adj studies).ab.  
 53 (data adj extraction).ab.  
 54 meta-analysis/  
 55 meta-analysis.ti.  
 56 comment.pt.  
 57 letter.pt.  
 58 editorial.pt.  
 59 animal/  
 60 human/  
**61 59 not (59 and 60)**  
**62 (or/49-55) not (56 or 57 or 58 or 61)**

**e) Selecting all reviews on diet and exercise modification in cancer survival**

**63 15 and 62 and (38 or 48)**

2) The second search strategy aims to identify all relevant epidemiological studies, which have been published since the last review. Since the date of the last review will be different for each type of cancer it is necessary to create individual search strategies for each cancer site. The searches will be limited according to the date of the last review. If no review exists on a particular cancer site then the databases will be searched from all years.

The search strategy listed below is for breast cancer. However, similar strategies will be created for each cancer site and adapted for use in all other databases.

### **Search strategy for BREAST CANCER in MEDLINE**

#### **a) Searching for all studies relating to breast cancer and survival**

- 1 exp Breast Neoplasms/
- 2 breast\$ adj3 neoplasm\$.tw
- 3 ((breast\$ adj3 cancer\$) or (breast\$ adj3 carcino\$)).tw.
- 4 or/1-3**
- 5 survivors/
- 6 exp survival analysis/
- 7 surviv\$.tw.
- 8 recurrence/
- 9 recur\$.tw.
- 10 quality of life/ or quality of life.tw. or qaly\$.tw.
- 11 mortality/
- 12 survival rate/
- 13 (manag\$ adj3 cancer\$).tw
- 14 or/5-13**
- 15 4 and 14**

#### **b) Searching for all studies relating to dietary modification**

- 16 exp "food and beverages"/
- 17 food\$.tw.
- 18 supplement\$.tw.
- 19 exp diet/
- 20 exp diet therapy/
- 21 diet\$.tw.
- 22 exp nutrition/
- 23 nutri\$.tw.
- 24 exp dietary fats/
- 25 exp dietary proteins/
- 26 exp dietary carbohydrates/
- 27 exp vitamins/
- 28 exp Feeding Behavior/
- 29 exp Drinking Behavior/
- 30 exp trace elements/
- 31 exp antioxidants/
- 32 exp Micronutrients/
- 33 calcium, dietary/
- 34 phosphorus, dietary/
- 35 exp sodium, dietary/

- 36 potassium, dietary/
- 37 iron, dietary/
- 38 or/16-37**

**c) Searching for all studies relating to exercise modification**

- 39 exp Exercise Movement Techniques/
- 40 exp exertion/
- 41 exp sports/
- 42 exp physical fitness/
- 43 exp exercise/
- 44 exercis\$.tw.
- 45 (physical\$ adj3 activ\$).tw.
- 46 (physical\$ adj3 fit\$).tw.
- 47 (physical\$ adj3 train\$).tw.
- 48 or/39-47**

**d) Filter for selecting only observational studies: case-control, cohort and cross-sectional studies (Scottish Intercollegiate Guidelines Network, June 2003).**

- 49 epidemiologic studies/
- 50 exp case control studies/
- 51 case control.tw
- 52 exp cohort studies/
- 53 (cohort adj (study or studies)).tw
- 54 cohort analy\$.tw
- 55 (Follow up adj (study or studies)).tw
- 56 (observational adj (study or studies)).tw
- 57 longitudinal.tw
- 58 retrospective.tw
- 59 exp cross-sectional studies/
- 60 cross sectional.tw
- 61 or/49-60**

**e) Selecting all epidemiological studies on diet and exercise modification in breast cancer survival**

- 62 15 and 61 and (38 or 48)**

### **3. Methods of the review**

#### **Data Collection**

Where possible the most up-to-date review will be used to obtain initial data for each cancer site.

Inclusion (or otherwise) of epidemiological studies published since the latest review will be assessed using an in/out form. The inclusion of studies will be assessed independently by two assessors and differences between reviewer's results will be resolved by discussion and, when necessary, in consultation with a third reviewer.

Data abstraction forms will be specifically designed. Data abstraction will be performed independently by two researchers and differences between reviewer's results will be resolved by returning to the relevant literature, discussion and, when necessary, consultation with a third reviewer.



## **Data Analysis**

Observational studies, unlike RCTs, are susceptible to confounding and selection bias: as a result the findings from observational studies may be distorted. Simply increasing study numbers in observational studies will not necessarily reduce bias and confounding. In fact, smaller studies may be able to better characterise potential confounding factors and, therefore, give a more accurate picture than larger observational studies. For these reasons the general methods of meta-analysis – combining data and weighting studies according to their statistical size (the larger the study the greater the weight) – have been suggested inappropriate with regards observational studies (Egger et al, 2001). Meta-analysis of observational studies will, therefore, not be undertaken in this review.

However, whilst meta-analysis is inappropriate a systematic review of the data is both plausible and appropriate. Where possible analysis will first focus on converting the results from each study into a suitable standard format to allow easy comparisons of results between studies. These results will then be tabulated and where appropriate ‘Forest Plots’ generated to give a graphical display of comparable results, together with their 95% confidence intervals, from each study. Any heterogeneity between study findings will then be investigated using sensitivity analysis to test the stability of findings across different study designs, approaches to exposure ascertainment and to selection of study participants.

Using these methods, as outlined by Egger and colleagues (Egger et al, 2001), we will avoid generating spurious results and misleading conclusions which are otherwise a common problem when reviewing data from a number of observational studies which differ in study design. Furthermore, it will assist considerably in deciding what future RCTs may be usefully initiated.

As already stated above confounding is an important consideration in observational studies and as part of the data collection we will note how confounding has been dealt with.

All analysis will be undertaken using the statistical package STATA, version 8.

## References

AICR. Dietary Options for Cancer Survivors. Washington: American Institute of Cancer Research; 2002a

AICR. Nutrition After Cancer. Washington: American Institute of Cancer Research; 2002b

Centre for Reviews and Dissemination (CRD). Undertaking Systematic Reviews of Research on Effectiveness. CRD's Guidance for those Carrying Out or Commissioning Reviews. CRD Report No. 4. 2<sup>nd</sup> ed. York: CRD; 2001.

Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994; 309:1286-91.

Egger M, Davey Smith G, Schneider M. Systematic reviews of observational studies. In: Egger M, Davey Smith G, Altman DG, eds. *Systematic Reviews in Health Care. Meta-analysis in context*. London: BMJ Books; 2001. p. 211-227.

Higgins JPT and Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539-1558.

Scottish Intercollegiate Guidelines Network (SIGN). Search Filters. Available from: <http://www.sign.ac.uk/methodology/filters.html#obs>. Accessed June 23, 2003.

Sterne JAC, Gavaghan D, Egger M (2000). Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *J Clin Epidemiol* 2000;53:1119-1129.

Sterne JAC, Egger M, Davey Smith G. Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001a ;323:101-105.

Sterne JAC, Bradburn MJ, Egger M. Meta-analysis in Stata. In: Egger M, Davey Smith G, Altman DG, eds. *Systematic Reviews in Health Care. Meta-analysis in context*. London: BMJ Books; 2001b. p. 347-69.

WHO. The World Health Report. Geneva: WHO; 1997.

WCRF and AICR. Food, Nutrition and the Prevention of Cancer: a Global Perspective. Washington: American Institute of Cancer Research; 1997.

**Appendix 1**

**In/Out Form**

Diet, exercise and cancer survival – review of RCTs.

**Identification details:**

Author: .....

Year: .....

Journal Reference: .....

On RefMan database? .....

**Study Selection Criteria (please circle):**

1. Study design is an RCT, not before/after or other design? Y/N

2. The study concerns individuals diagnosed with cancer Y/N  
• The study can include men and women of any age

3. The study concerns a dietary modification and/or an exercise intervention Y/N

- Dietary modification includes any life-style dietary modification, dietary education, nutrient supplementation or complementary medicine diet-based therapy.
- There is no restriction on route of dietary intake used.
- Diet and life-style modifications consequent on the disease or its treatment will NOT be included. These are peri- and post-operative dietary modification together with calorie enhancement for cancer cachexia.
- All exercise interventions are included.
- Modifications and interventions are included regardless of their time period.

**Please cross only one box:**

**In**

**Out**

**Pending**

**Appendix 2**

**DIET, EXERCISE AND CANCER SURVIVAL:**  
**RCT DATA EXTRACTION FORM**

(EDITION 15.7.03.)

**1. ID, REFS AND DATA SOURCES**

**TRIAL ID NO** .....

**COMPARISON NO**..... (placebo/usual care)

**REVIEWER** .....

**DATE** .....

**STUDY NAME OR ACRONYM IF ANY**.....

*Source reference:*

Database Reference ID number

FIRST AUTHOR .....

YEAR.....

If other references for this trial, list Reference ID numbers

**OTHER ID NOS FREE TEXT** ... ..

.....

.....

**PUBLICATION TYPE (MOST ACCESSIBLE) (please circle):**

Full Journal Article/Journal Abstract

**ANALYSABLE DATA CAN ONLY BE OBTAINED VIA PERSONAL  
COMMUNICATION? (please circle):**

YES/NO

**2. OVERALL DESIGN DETAILS**

**TARGET GROUP (please circle all that apply):**

Neonates/Children pre-teens/Teens/ Children+teens/Young Adults/  
Middle-Aged/Elderly/Working Age Adults/All Adults.

**TARGET GROUP FREE TEXT (e.g. adults over 40, children under 5):**

.....

**WHAT ARE THE INCLUSION AND EXCLUSION CRITERIA FREETEXT:**

.....

.....

**CANCER TYPE**

Breast / mouth / pharynx / nasopharynx / larynx / oesophagus / lung / stomach /  
pancreas / gallbladder / liver / colon / rectum / ovary / endometrium / cervix / thyroid /  
kidney / bladder / prostate / skin / bone / nervous tissue / haematological / lymphatic

**OTHER CANCER TYPE FREETEXT .....**

.....

**LOCATION OF POPULATION FREETEXT (e.g. all patients registered at a GP in  
the UK who are diagnosed with cancer, all out-patients attending a cancer clinic at a  
US hospital): .....**

.....

.....

**EXPERIMENTAL INTERVENTION WAS:**

Exercise intervention / Life-style dietary change / dietary supplementation / food-  
based complementary medicine therapy /Individual Dietary Counselling/Group  
Dietary Education/ Dietary information with minimum personal contact **(NB: select  
all that apply)**

**OTHER INTERVENTION FREETEXT .....**

.....

**2. OVERALL DESIGN DETAILS (CONTINUED)**

TOTAL PERIOD OF ACTIVE INTERVENTION COVERED BY THIS REPORT  
(please leave blank if not reported):

**Mean**.....

**Medium** .....

**Range**.....

TYPE OF CONTROL CONDITION WAS:

Placebo Personal Contact/ No-intervention-usual diet/ No intervention – usual exercise / General health education materials/ Placebo supplement  
**(NB: select all that apply)**

**HOW WAS COMPLIANCE CHECKED?**

Monitoring of consumption by researcher/ self-report / other’s report on subject (e.g. relative) / weighed food intake diary/ <7 days dietary diary/ 7+ days dietary diary/ biochemical levels in the blood or urine / food frequency questionnaire / physical activity monitoring machine / 24-hour exercise diary / <7 days exercise diary / 7+ days exercise diary / exercise frequency questionnaire / it was not checked  
**(NB select all that apply)**

HOW MANY INTERVENTION ARMS WERE THERE IN THIS STUDY?

Number = .....

**DETAILS OF THE INTERVENTIONS IN THIS STUDY FREETEXT:**

.....  
.....  
.....

FREETEXT COMMENT ON DESIGN e.g. DIETARY / EXERCISE MEASUREMENTS:

.....  
.....

**3. DETAILS GIVEN OF PARTICIPANTS IN THIS COMPARISON**

	<b>Experimental Arm</b>	<b>Control Arm</b>
No. Assigned (to each arm)		
Percentage of Males		
Mean Age		
Median age		
Percentage Non-White		
No. judged non-compliant (failed to complete the intervention)		
No. Analysed for outcome measure		
• Survival / All-cause mortality		
• Cancer mortality		
• Primary cancer recurrence		
• Second primary cancer		
• Quality of life		
• Side effects		

NB - an expanded version of this table will be used for studies with more than one intervention arm.

**WERE THE DETAILS GIVEN (AS ABOVE) OF PARTICIPANTS FOR (please circle):**

The assigned/ the analysed/ both

**FREETEXT ON FLOW OF PEOPLE THROUGH TRIAL INCLUDING ANY SWITCHES IN TREATMENT CONDITION**

.....

.....

.....

**4. OUTCOMES USED AND ANALYSIS**

**WHAT OUTCOME MEASURES WERE USED:**

Survival or all-cause mortality / cancer mortality / primary cancer recurrence/ second primary cancer/ quality of life / side effects

**WERE OUTCOMES ANALYSED ON AN INTENTION TO TREAT BASIS**

Survival/all-cause mortality ..... Yes /No /Not clear  
cancer mortality ..... Yes /No /Not clear  
primary cancer recurrence ..... Yes /No /Not clear  
second primary cancer..... Yes /No /Not clear  
quality of life ..... Yes /No /Not clear  
side effects..... Yes /No /Not clear

**FREETEXT ON OUTCOMES USED AND ANALYSIS:**

.....  
.....  
.....  
.....



**5. OUTCOME DATA**

	<b>Intervention Events</b>	<b>Intervention group size</b>	<b>Control Events</b>	<b>Control group size</b>
Denominators: persons				
Denominators: person-years				
No. deaths from all -cause mortality*				
No. deaths from cancer*				
No. with a primary cancer recurrence*				
No. with a second primary cancer recurrence*				

\* Extract maximum of one event per person per outcome, if more than one event occurs per outcome then only the first one will be recorded

**FREETEXT COMMENT ON QUALITY OF LIFE OUTCOME MEASUREMENTS:**

.....

.....

.....

.....

**FREETEXT COMMENT ON SIDE EFFECTS OUTCOME MEASUREMENTS:**

.....

.....

.....

.....

**6. QUALITY**

**SELECTION BIAS**

Method of Randomisation: ALLOCATION COMPUTER GENERATED/CENTRAL OFFICE-TELEPHONE SERVICE/SEALED ENVELOPES/OTHER/NOT CLEAR

other (specify) **FREETEXT**.....

**PERFORMANCE AND DETECTION BIAS**

Were participants blind to the intervention assigned? YES/NO/NOT CLEAR

Were the clinicians/investigators blind to which intervention was being provided? YES/NO/NOT CLEAR

Were the assessors of outcome measures blind to the intervention provided? YES/NO/NOT CLEAR

Were the statisticians blind to the intervention provided? YES/NO/NOT CLEAR

**LOSSES TO FOLLOW UP**

**How many subjects were lost to outcome follow up?**

- Survival / all-cause mortality.....
- cancer mortality.....
- primary cancer recurrence .....
- second primary cancer .....
- quality of life .....
- side effects .....

**Are subjects lost to follow up enumerated with reasons for loss?**

Survival / all-cause mortality ..... Yes /No /Not clear  
cancer mortality ..... Yes /No /Not clear  
primary cancer recurrence ..... Yes /No /Not clear  
second primary cancer..... Yes /No /Not clear  
quality of life ..... Yes /No /Not clear  
side effects..... Yes /No /Not clear

**Are they included in the denominator for analyses of outcomes?**

Survival / all-cause mortality ..... Yes /No /Not clear  
cancer mortality ..... Yes /No /Not clear  
primary cancer recurrence ..... Yes /No /Not clear  
second primary cancer..... Yes /No /Not clear  
quality of life ..... Yes /No /Not clear  
side effects..... Yes /No /Not clear

**How were outcome measurements adjusted for losses to follow up?**

Survival / all-cause mortality ..... Yes /No /Not clear  
cancer mortality ..... Yes /No /Not clear  
primary cancer recurrence ..... Yes /No /Not clear  
second primary cancer..... Yes /No /Not clear  
quality of life ..... Yes /No /Not clear  
side effects..... Yes /No /Not clear

END OF EXTRACTION SCHEDULE

*Continuous update of the WCRF-AICR report on diet and cancer*

Protocol: Breast cancer survivors

Prepared by: Imperial College Team

Search strategy

We will use the standard search strategy developed by WCRF. The original search strategy was adapted for OVID and EMBASE by a medical librarian.

**Breast Cancer Survival Search Strategy (OVID Medline)**

1. cancer\$.ab,ti.
2. neoplasm\$.ab,ti.
3. tumour\$.ab,ti.
4. tumor\$.ab,ti.
5. carcinoma\$.ab,ti.
6. adenocarcinoma\$.ab,ti.
7. 1 or 2 or 3 or 4 or 5 or 6
8. mammary.ab,ti.
9. breast.ab,ti.
10. 8 or 9
- 11. 7 and 10**
12. treatment\$.ab,ti.
13. (second\$ adj5 primar\$).ab,ti.
14. remission\$.ab,ti.
15. Treatment Outcome/
16. Disease-Free Survival/
17. Remission Induction/
18. Survivors/
19. Survival Analysis/
20. surviv\$.ab,ti.
21. recur\$.ab,ti.
22. "quality of life".ab,ti.
23. qaly\$.ab,ti.
24. Mortality/

25. Survival Rate/
26. (cancer\$ adj5 manag\$).ab,ti.
27. (cancer\$ adj5 treat\$).ab,ti.
28. rehabilitation.ab,ti.
- 29. 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28**
30. Diet Therapy/
31. nutrition.ab,ti.
32. (diet or diets or dietic or dietary or eating or intake or nutrient\$ or nutrition or vegetarian\$ or vegan\$ or (seventh adj1 day adj1 adventist) or macrobiotic or breastfeed\$ or breast feed\$ or breastfed or breast fed or breastmilk or breast milk).ab,ti.
- 33. 30 or 31 or 32**
34. Food/
35. (food\$ or cereal\$ or grain\$ or granary or wholegrain or wholewheat or roots or plantain\$ or tuber or tubers or vegetable\$ or fruit\$ or pulses or beans or lentils or chickpeas or legume\$ or soy or soya or nut or nuts or peanut\$ or groundnut\$ or seeds).ab,ti.
36. (meat or beef or pork or lamb or poultry or chicken or turkey or duck or fish or fat or fats or fatty or egg or eggs or bread or oils or shellfish or seafood or sugar or syrup or dairy or milk or herbs or spices or chilli or chillis or pepper\$ or condiments).ab,ti.
- 37. 34 or 35 or 36**
38. Beverages/
39. (fluid intake or water or drinks or drinking or tea or coffee or caffeine or juice or beer or spirits or liquor or wine or alcohol or alcoholic or beverage\$ or ethanol or yerba mate or ilex or paraguariensis).ab,ti.
- 40. 38 or 39**
41. Pesticides/
42. Fertilizers/
43. Veterinary Drugs/
44. (pesticide\$ or herbicide\$ or DDT or fertiliser\$ or fertilizer\$ or organic or contaminants or contaminate\$ or veterinary drug\$ or polychlorinated dibenzofuran\$ or PCDF\$ or polychlorinated dibenzodioxin\$ or PCDD\$ or polychlorinated biphenyl\$ or PCB\$ or cadmium or arsenic or chlorinated hydrocarbon\$ or microbial contamination\$).ab,ti.
- 45. 41 or 42 or 43 or 44**
46. Food Preservation/
47. (mycotoxin\$ or aflatoxin\$ or pickled or bottled or bottling or canned or canning or vacuum pack\$ or refrigerate\$ or refrigeration or cured or smoked or preserved or

preservatives or nitrosamine or hydrogenation or fortified or additive\$ or colouring\$ or coloring\$ or flavouring\$ or flavoring\$ or nitrates or nitrites or solvent or solvents or ferment\$ or processed or antioxidant\$ or genetic modif\$ or genetically modif\$ or vinyl chloride or packaging or labelling or phthalates).ab,ti.

**48. 46 or 47**

49. Cooking/

50. (cooking or cooked or grill or grilled or fried or fry or roast or bake or baked or stewing or stewed or casserol\$ or broil or broiled or boiled or microwave or microwaved or reheating or reheating or heating or re-heated or heated or poach or poached or steamed or barbecue\$ or chargrill\$ or heterocyclic amines or polycyclic aromatic hydrocarbons).ab,ti.

**51. 49 or 50**

52. Dietary Carbohydrates/

53. Dietary Proteins/

54. Sweetening Agents/

55. (salt or salting or salted or fiber or fibre or polysaccharide\$ or starch or starchy or carbohydrate\$ or lipid\$ or linoleic acid\$ or sterols or stanols or sugar\$ or sweetener\$ or saccharin\$ or aspartame or acesulfame or cyclamates or maltose or mannitol or sorbitol or sucrose or xylitol or cholesterol or hydrogenated dietary oils or hydrogenated lard or hydrogenated oils or dietary protein or dietary proteins or protein intake or protein consumption or total protein\$ or animal protein\$ or vegetable protein\$ or plant protein\$).ab,ti.

**56. 52 or 53 or 54 or 55**

57. Vitamins/

58. (supplements or supplement or vitamin\$ or retinol or carotenoid\$ or tocopherol or folate\$ or folic acid or methionine or riboflavin or thiamine or niacin or pyridoxine or cobalamin or mineral\$ or sodium or iron or calcium or selenium or iodine or magnesium or potassium or zinc or copper or phosphorus or manganese or chromium or phytochemical or allium or isothiocyanate\$ or glucosinolate\$ or indoles or polyphenol\$ or phytoestrogen\$ or genistein or saponin\$ or coumarin\$).ab,ti.

**59. 57 or 58**

60. Physical Fitness/

61. Physical Exertion/

62. Physical Endurance/

63. Walking/

64. Exercise Movement Techniques/

- 65. Sports/
- 66. Exercise/
- 67. Dancing/
- 68. (recreational activit\$ or household activit\$ or occupational activit\$ or physical activit\$ or physical inactivit\$ or exercise\$ or exercising or energy intake or energy expenditure or energy balance or energy density).ab,ti.

**69. 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68**

- 70. Growth/
- 71. Anthropometry/
- 72. Body Composition/
- 73. Body Constitution/
- 74. (weight loss or weight gain or anthropometry or birth weight or birthweight or birth-weight or child development or height or body composition or body mass or BMI or obesity or obese or overweight or over-weight or over weight or skinfold measurement\$ or skinfold thickness or DEXA or bio-impedence or waist circumference or hip circumference or waist hip ratio\$).ab,ti.

**75. 70 or 71 or 72 or 73 or 74**

- 76. Plants/
- 77. Plant Preparations/
- 78. (psk\$ or krestin\$ or (retinoic\$ adj acid\$) or isotretino\$ or tretino\$).ab,ti.
- 79. Retinoids/

**80. 76 or 77 or 78 or 79**

- 81. Holistic Health/
- 82. Homeopathy/
- 83. Medicine, Traditional/
- 84. Medicine, African Traditional/
- 85. Medicine, Arabic/
- 86. Medicine, Ayurvedic/
- 87. Medicine, East Asian Traditional/
- 88. Medicine, Chinese Traditional/
- 89. Medicine, Kampo/
- 90. Medicine, Mongolian Traditional/
- 91. Tai Ji/
- 92. Yoga/
- 93. Naturopathy/
- 94. Phytotherapy/

- 95. Homeopathy/
- 96. Complementary Therapies/
- 97. (nutraceutical\$ or neutraceutical\$ or alternative medicine or complementary medicine or alternative therap\$).ab,ti.
- 98. 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97**
- 99. 33 or 37 or 40 or 45 or 48 or 51 or 56 or 59 or 69 or 75 or 80 or 98
- 100. 11 and 29 and 99
- 101. exp animal/
- 102. exp human/
- 103. 101 not 102
- 104. 100 not 103

### **Breast Cancer Survival Search Strategy (OVID Embase)**

- 1. cancer\$.ab,ti.
- 2. neoplasm\$.ab,ti.
- 3. tumour\$.ab,ti.
- 4. tumor\$.ab,ti.
- 5. carcinoma\$.ab,ti.
- 6. adenocarcinoma\$.ab,ti.
- 7. 1 or 2 or 3 or 4 or 5 or 6
- 8. mammary.ab,ti.
- 9. breast.ab,ti.
- 10. 8 or 9
- 11. 7 and 10**
- 12. treatment\$.ab,ti.
- 13. (second\$ adj5 primar\$).ab,ti.
- 14. remission\$.ab,ti.
- 15. Treatment Outcome/
- 16. Disease-Free Survival/
- 17. remission induction.ti,ab.
- 18. Survivor/
- 19. Survival Analysis/



20. surviv\$.ab,ti.
21. recur\$.ab,ti.
22. "quality of life".ab,ti.
23. qaly\$.ab,ti.
24. Mortality/
25. Survival Rate/
26. (cancer\$ adj5 manag\$).ab,ti.
27. (cancer\$ adj5 treat\$).ab,ti.
28. rehabilitation.ab,ti.
- 29. 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28**
30. Diet Therapy/
31. Nutrition/
32. (diet or diets or dietic or dietary or eating or intake or nutrient\$ or nutrition or vegetarian\$ or vegan\$ or (seventh adj1 day adj1 adventist) or macrobiotic or breastfeed\$ or breast feed\$ or breastfed or breast fed or breastmilk or breast milk).ab,ti.
- 33. 30 or 31 or 32**
34. Food/
35. (food\$ or cereal\$ or grain\$ or granary or wholegrain or wholewheat or roots or plantain\$ or tuber or tubers or vegetable\$ or fruit\$ or pulses or beans or lentils or chickpeas or legume\$ or soy or soya or nut or nuts or peanut\$ or groundnut\$ or seeds).ab,ti.
36. (meat or beef or pork or lamb or poultry or chicken or turkey or duck or fish or fat or fats or fatty or egg or eggs or bread or oils or shellfish or seafood or sugar or syrup or dairy or milk or herbs or spices or chilli or chillis or pepper\$ or condiments).ab,ti.
- 37. 34 or 35 or 36**
38. Beverages/
39. (fluid intake or water or drinks or drinking or tea or coffee or caffeine or juice or beer or spirits or liquor or wine or alcohol or alcoholic or beverage\$ or ethanol or yerba mate or ilex or paraguariensis).ab,ti.
- 40. 38 or 39**
41. Pesticides/
42. Fertilizers/
43. Veterinary Drugs/
44. (pesticide\$ or herbicide\$ or DDT or fertiliser\$ or fertilizer\$ or organic or contaminants or contaminate\$ or veterinary drug\$ or polychlorinated dibenzofuran\$ or PCDF\$ or

polychlorinated dibenzodioxin\$ or PCDD\$ or polychlorinated biphenyl\$ or PCB\$ or cadmium or arsenic or chlorinated hydrocarbon\$ or microbial contamination\$).ab,ti.

**45. 41 or 42 or 43 or 44**

46. Food Preservation/

47. (mycotoxin\$ or aflatoxin\$ or pickled or bottled or bottling or canned or canning or vacuum pack\$ or refrigerate\$ or refrigeration or cured or smoked or preserved or preservatives or nitrosamine or hydrogenation or fortified or additive\$ or colouring\$ or coloring\$ or flavouring\$ or flavoring\$ or nitrates or nitrites or solvent or solvents or ferment\$ or processed or antioxidant\$ or genetic modif\$ or genetically modif\$ or vinyl chloride or packaging or labelling or phthalates).ab,ti.

**48. 46 or 47**

49. Cooking/

50. (cooking or cooked or grill or grilled or fried or fry or roast or bake or baked or stewing or stewed or casserol\$ or broil or broiled or boiled or microwave or microwaved or reheating or reheating or heating or re-heated or heated or poach or poached or steamed or barbecue\$ or chargrill\$ or heterocyclic amines or polycyclic aromatic hydrocarbons).ab,ti.

**51. 49 or 50**

52. Dietary carbohydrate\$.ti,ab.

53. protein intake/

54. Sweetening Agents/

55. (salt or salting or salted or fiber or fibre or polysaccharide\$ or starch or starchy or carbohydrate\$ or lipid\$ or linoleic acid\$ or sterols or stanols or sugar\$ or sweetener\$ or saccharin\$ or aspartame or acesulfame or cyclamates or maltose or mannitol or sorbitol or sucrose or xylitol or cholesterol or hydrogenated dietary oils or hydrogenated lard or hydrogenated oils or dietary protein or dietary proteins or protein intake or protein consumption or total protein\$ or animal protein\$ or vegetable protein\$ or plant protein\$).ab,ti.

**56. 52 or 53 or 54 or 55**

57. Vitamins/

58. (supplements or supplement or vitamin\$ or retinol or carotenoid\$ or tocopherol or folate\$ or folic acid or methionine or riboflavin or thiamine or niacin or pyridoxine or cobalamin or mineral\$ or sodium or iron or calcium or selenium or iodine or magnesium or potassium or zinc or copper or phosphorus or manganese or chromium or phytochemical or allium or isothiocyanate\$ or glucosinolate\$ or indoles or polyphenol\$ or phytoestrogen\$ or genistein or saponin\$ or coumarin\$).ab,ti.

**59. 57 or 58**

60. FITNESS/

61. physical exertion.ti,ab.

62. ENDURANCE/

63. Walking/

64. exercise movement technique\$.ti,ab.

65. SPORT/

66. Exercise/

67. Dancing/

68. (recreational activit\$ or household activit\$ or occupational activit\$ or physical activit\$ or physical inactivit\$ or exercise\$ or exercising or energy intake or energy expenditure or energy balance or energy density).ab,ti.

**69. 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68**

70. Growth/

71. Anthropometry/

72. Body Composition/

73. Body Constitution/

74. (weight loss or weight gain or anthropometry or birth weight or birthweight or birthweight or child development or height or body composition or body mass or BMI or obesity or obese or overweight or over-weight or over weight or skinfold measurement\$ or skinfold thickness or DEXA or bio-impedence or waist circumference or hip circumference or waist hip ratio\$).ab,ti.

**75. 70 or 71 or 72 or 73 or 74**

76. Plant/

77. plant medicinal product/

78. (psk\$ or krestin\$ or (retinoic\$ adj acid\$) or isotretino\$ or tretino\$).ab,ti.

79. Retinoids/

**80. 76 or 77 or 78 or 79**

81. holistic health.ti,ab.

82. Homeopathy/

83. traditional medicine/

84. African medicine/

85. arabic medicine.ti,ab.

86. ayurvedic medicine.ti,ab.

87. oriental medicine/

88. Chinese medicine/

- 89. mongolian traditional medicine.ti,ab.
- 90. Tai Chi/
- 91. Yoga/
- 92. naturopathy.ti,ab.
- 93. Phytotherapy/
- 94. Homeopathy/
- 95. complementary therap\$.ti,ab.
- 96. (nutraceutical\$ or neutraceutical\$ or alternative medicine or complementary medicine or alternative therap\$).ab,ti.
- 97. 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96**
- 98. 33 or 37 or 40 or 45 or 48 or 51 or 56 or 59 or 69 or 75 or 80 or 97
- 99. 11 and 29 and 98
- 100. exp animal/
- 101. exp human/
- 102. 100 not 101
- 103. 99 not 102

## Appendix 4. Exposures in WCRF database

### Main headings and sub-headings.

#### 1 Patterns of diet

##### 1.1 Regionally defined diets

*Include all regionally defined diets, evident in the literature. These are likely to include Mediterranean, Mesoamerican, oriental, including Japanese and Chinese, and “western type”.*

##### 1.2 Socio-economically defined diets

*To include diets of low-income, middle-income and high-income countries (presented, when available in this order). Rich and poor populations within low-income, middle-income and high-income countries should also be considered. This section should also include the concept of poverty diets (monotonous diets consumed by impoverished populations in the economically-developing world mostly made up of one starchy staple, and may be lacking in micronutrients).*

##### 1.3 Culturally defined diets

*To include dietary patterns such as vegetarianism, vegan diets, macrobiotic diets and diets of Seventh-day Adventists.*

##### 1.4 Individual level dietary patterns

*To include work on factor and cluster analysis, and various scores and indexes (e.g. Mediterranean type diet index) that do not fit into the headings above.*

##### 1.5 Other dietary patterns

*Include under this heading any other dietary patterns present in the literature, that are not regionally, socio-economically, culturally or individually defined.*

##### 1.6 Breastfeeding

###### 1.6.1 Mother

###### 1.6.2 Child

*Results concerning the effects of breastfeeding on the development of cancer should be disaggregated into effects on the mother and effects on the child. Wherever possible*

*detailed information on duration of total and exclusive breastfeeding, and of complementary feeding should be included.*

## 1.7 Other issues

*For example results related to meal frequency, frequency of snacking, dessert-eating and breakfast-eating should be reported here.*

## 2 Foods

### 2.1 Starchy foods

#### 2.1.1 Cereals (grains)

##### 2.1.1.1 Wholegrain cereals and cereal products

##### 2.1.1.2 Refined cereals and cereal products

#### 2.1.2 Starchy roots, tubers and plantains

#### 2.1.3 Other starchy foods

### 2.2 Fruit and (non-starchy) vegetables

*Results for “fruit and vegetables” should be reported here. If the definition of vegetables used here is different from that used in the first report, this should be highlighted.*

#### 2.2.1 Non-starchy vegetables

*This heading should be used to report total non-starchy vegetables. If results about specific vegetables are reported they should be recorded under one of the sub-headings below or if not covered, they should be recorded under ‘2.2.1.5 other’.*

##### 2.2.1.1 Non-starchy root vegetables and tubers

##### 2.2.1.2 Cruciferous vegetables

##### 2.2.1.3 Allium vegetables

##### 2.2.1.4 Green leafy vegetables (not including cruciferous vegetables)

##### 2.2.1.5 Other non-starchy vegetables

*Other non-starchy vegetables’ should include foods that are botanically fruits but are eaten as vegetables, e.g. tomatoes, courgettes. In addition vegetables such as French beans that do not fit into the other categories, above.*

*If there is another sub-category of vegetables that does not easily fit into a category above eg salted root vegetables (ie you do not know if it is starchy or not) then report under 2.2.1.5. and note the precise definition used by the study. Note that the eg salted root vegetables should also be reported under 4.2.5.3 salted foods. If in doubt, enter the exposure more than once in this way.*

#### 2.2.1.6 Raw vegetables

*This section should include any vegetables specified as eaten raw. Results concerning specific groups and type of raw vegetable should be reported twice i.e. also under the relevant headings 2.2.1.1 –2.2.1.5.*

#### 2.2.2 Fruits

##### 2.2.2.1 Citrus fruit

##### 2.2.2.2 Other

*If results are available that consider other groups of fruit or a particular fruit please report under 'other', specifying the grouping/fruit used in the literature.*

#### 2.3 Pulses (legumes)

*To include soya and soya products, peanuts (groundnuts), chickpeas, lentils. Where results are available for a specific pulse/legume, e.g. soya, please report under a separate heading.*

#### 2.4 Nuts and Seeds

*To include all tree nuts and seeds, but not peanuts (groundnuts). Where results are available for a specific nut/seed, e.g. brazil nuts, please report under a separate heading.*

#### 2.5 Meat, poultry, fish and eggs

*Wherever possible please differentiate between farmed and wild meat, poultry and fish.*

##### 2.5.1 Meat

*This heading refers only to red meat: essentially beef, lamb, pork from farmed domesticated animals either fresh or frozen, or dried without any other form of preservation. It does not refer to poultry or fish.*

*Where there are data for offal (organs and other non-flesh parts of meat) and also when there are data for wild and non-domesticated animals, please show these separately under this general heading as a subcategory.*

##### 2.5.1.1 Fresh Meat

##### 2.5.1.2 Processed meat

*Repeat results concerning processed meat here and under the relevant section under 4. Food Production and Processing. Please record the definition of 'processed meat' used by each study.*

### 2.5.1.3 Red meat

*Where results are available for a particular type of meat, e.g. beef, pork or lamb, please report under a separate heading.*

*Show any data on wild meat (game) under this heading as a separate sub-category.*

### 2.5.1.4 Poultry

*Show any data on wild birds under this heading as a separate sub-category.*

### 2.5.2 Fish

*Wherever results are available for particular types of fish e.g. oily fish, white fish, please report under separate headings.*

### 2.5.3 Shellfish and other seafood

### 2.5.4 Eggs

## 2.6 Fats, oils and sugars

### 2.6.1 Animal fats

### 2.6.2 Plant oils

### 2.6.3 Hydrogenated fats and oils

*Results concerning hydrogenated fats and oils should be reported twice, here and under 4.3.2 Hydrogenation*

### 2.6.4 Sugars

*This heading refers to added (extrinsic) sugars and syrups as a food, that is refined sugars, such as table sugar, or sugar used in bakery products.*

## 2.7 Milk and dairy products

*Results concerning milk should be reported twice, here and under 3.3 Milk*

## 2.8 Herbs, spices, condiments

*The 1997 report found evidence concerning turmeric, saffron, cumin, ginger, pepper, chilli pepper and harissa.*

## 2.9 Composite foods



*Eg, snacks, crisps, desserts, pizza. Also report any mixed food exposures here ie if an exposure is reported as a combination of 2 or more foods that cross categories (eg bacon and eggs). Label each mixed food exposure.*

### 3 Beverages

#### 3.1 Total fluid intake

#### 3.2 Water

#### 3.3 Milk

*For results concerning milk please report twice, here and under 2.7 Milk and Dairy Products.*

#### 3.4 Soft drinks

*Soft drinks that are both carbonated and sugary should be reported under this general heading. Drinks that contain artificial sweeteners should be reported separately and labelled as such.*

##### 3.4.1 Sugary (not carbonated)

##### 3.4.2 Carbonated (not sugary)

*The precise definition used by the studies should be highlighted, as definitions used for various soft drinks vary greatly.*

#### 3.5 Fruit juices

*The precise definition used by the studies should be highlighted, as definitions used for various fruit juices vary greatly.*

#### 3.6 Hot drinks

##### 3.6.1 Coffee

##### 3.6.2 Tea

*Report herbal tea as a sub-category under tea.*

##### 3.6.2.1 Black tea

##### 3.6.2.2 Green tea

##### 3.6.3 Maté

##### 3.6.4 Other hot drinks

#### 3.7 Alcoholic drinks

##### 3.7.1 Total

- 3.7.1.1 Beers
- 3.7.1.2 Wines
- 3.7.1.3 Spirits
- 3.7.1.4 Other alcoholic drinks

## 4 Food production, preservation, processing and preparation

### 4.1 Production

- 4.1.1 Traditional methods (*to include 'organic'*)
- 4.1.2 Chemical contaminants

*Only results based on human evidence should be reported here (see instructions for dealing with mechanistic studies). Please be comprehensive and cover the exposures listed below:*

- 4.1.2.1 Pesticides
- 4.1.2.2 DDT
- 4.1.2.3 Herbicides
- 4.1.2.4 Fertilisers
- 4.1.2.5 Veterinary drugs
- 4.1.2.6 Other chemicals
  - 4.1.2.6.1 Polychlorinated dibenzofurans (PCDFs)
  - 4.1.2.6.2 Polychlorinated dibenzodioxins (PCDDs)
  - 4.1.2.6.3 Polychlorinated biphenyls (PCBs)
- 4.1.2.7 Heavy metals
  - 4.1.2.7.1 Cadmium
  - 4.1.2.7.2 Arsenic
- 4.1.2.8 Waterborne residues
  - 4.1.2.8.1 Chlorinated hydrocarbons
- 4.1.2.9 Other contaminants

*Please also report any results that cover the cumulative effect of low doses of contaminants in this section.*

### 4.2 Preservation

- 4.2.1 Drying
- 4.2.2 Storage
  - 4.2.2.1 Mycotoxins

- 4.2.2.1.1 Aflatoxins
- 4.2.2.1.2 Others
  
- 4.2.3 Bottling, canning, vacuum packing
- 4.2.4 Refrigeration
- 4.2.5 Salt, salting
  - 4.2.5.1 Salt
  - 4.2.5.2 Salting
  - 4.2.5.3 Salted foods
    - 4.2.5.3.1 Salted animal food
    - 4.2.5.3.2 Salted plant food
  
- 4.2.6 Pickling
- 4.2.7 Curing and smoking
  - 4.2.7.1 Cured foods
    - 4.2.7.1.1 Cured meats
    - 4.2.7.1.2 Smoked foods

*For some cancers e.g. colon, rectum, stomach and pancreas, it may be important to report results about specific cured foods, cured meats and smoked meats. N-nitrosamines should also be covered here.*

### 4.3 Processing

#### 4.3.1 Refining

*Results concerning refined cereals and cereal products should be reported twice, here and under 2.1.1.2 refined cereals and cereal products.*

#### 4.3.2 Hydrogenation

*Results concerning hydrogenated fats and oils should be reported twice, here and under 2.6.3 Hydrogenated fats and oils*

#### 4.3.3 Fermenting

#### 4.3.4 Compositional manipulation

##### 4.3.4.1 Fortification

##### 4.3.4.2 Genetic modification

##### 4.3.4.3 Other methods

#### 4.3.5 Food additives

##### 4.3.5.1 Flavours

*Report results for monosodium glutamate as a separate category under 4.3.5.1 Flavours.*

4.3.5.2 Sweeteners (non-caloric)

4.3.5.3 Colours

4.3.5.4 Preservatives

4.3.5.4.1 Nitrites and nitrates

4.3.5.5 Solvents

4.3.5.6 Fat substitutes

4.3.5.7 Other food additives

*Please also report any results that cover the cumulative effect of low doses of additives. Please also report any results that cover synthetic antioxidants*

4.3.6 Packaging

4.3.6.1 Vinyl chloride

4.3.6.2 Phthalates

4.4 Preparation

4.4.1 Fresh food

4.4.1.1 Raw

*Report results regarding all raw food other than fruit and vegetables here. There is a separate heading for raw fruit and vegetables (2.2.1.6).*

4.4.1.2 Juiced

4.4.2 Cooked food

4.4.2.1 Steaming, boiling, poaching

4.4.2.2 Stewing, casseroling

4.4.2.3 Baking, roasting

4.4.2.4 Microwaving

4.4.2.5 Frying

4.4.2.6 Grilling (broiling) and barbecuing

4.4.2.7 Heating, re-heating

*Some studies may have reported methods of cooking in terms of temperature or cooking medium, and also some studies may have indicated whether the food was cooked in a direct or indirect flame. When this information is available, it should be included in the SLR report.*

*Results linked to mechanisms e.g. heterocyclic amines, acrylamides and polycyclic aromatic hydrocarbons should also be reported here. There may also be some literature on burned food that should be reported in this section.*

## 5 Dietary constituents

*Food constituents' relationship to outcome needs to be considered in relation to dose and form including use in fortified foods, food supplements, nutrient supplements and specially formulated foods. Where relevant and possible these should be disaggregated.*

### 5.1 Carbohydrate

#### 5.1.1 Total carbohydrate

#### 5.1.2 Non-starch polysaccharides/dietary fibre

##### 5.1.2.1 Cereal fibre

##### 5.1.2.2 Vegetable fibre

##### 5.1.2.3 Fruit fibre

#### 5.1.3 Starch

##### 5.1.3.1 Resistant starch

#### 5.1.4 Sugars

*This heading refers to intrinsic sugars that are naturally incorporated into the cellular structure of foods, and also extrinsic sugars not incorporated into the cellular structure of foods. Results for intrinsic and extrinsic sugars should be presented separately. Count honey and sugars in fruit juices as extrinsic. They can be natural and unprocessed, such as honey, or refined such as table sugar. Any results related to specific sugars e.g. fructose should be reported here.*

### 5.2 Lipids

#### 5.2.1 Total fat

#### 5.2.2 Saturated fatty acids

#### 5.2.3 Monounsaturated fatty acids

#### 5.2.4 Polyunsaturated fatty acids

##### 5.2.4.1 n-3 fatty acids

*Where available, results concerning alpha linolenic acid and long chain n-3 PUFA should be reported here, and if possible separately.*

##### 5.2.4.2 n-6 fatty acids

##### 5.2.4.3 Conjugated linoleic acid

#### 5.2.5 Trans fatty acids

#### 5.2.6 Other dietary lipids, cholesterol, plant sterols and stanols.

*For certain cancers, e.g. endometrium, lung, and pancreas, results concerning dietary cholesterol may be available. These results should be reported under this section.*

### 5.3 Protein

- 5.3.1 Total protein
- 5.3.2 Plant protein
- 5.3.3 Animal protein

### 5.4 Alcohol

*This section refers to ethanol the chemical. Results related to specific alcoholic drinks should be reported under 3.7 Alcoholic drinks.*

### 5.5 Vitamins

#### 5.5.1 Vitamin A

- 5.5.1.1 Retinol
- 5.5.1.2 Provitamin A carotenoids

#### 5.5.2 Non-provitamin A carotenoids

*Record total carotenoids under 5.5.2 as a separate category marked Total Carotenoids.*

#### 5.5.3 Folates and associated compounds

*Examples of the associated compounds are lipotropes, methionine and other methyl donors.*

- 5.5.4 Riboflavin
- 5.5.5 Thiamin (vitamin B1)
- 5.5.6 Niacin
- 5.5.7 Pyridoxine (vitamin B6)
- 5.5.8 Cobalamin (vitamin B12)
- 5.5.9 Vitamin C
- 5.5.10 Vitamin D (and calcium)
- 5.5.11 Vitamin E
- 5.5.12 Vitamin K
- 5.5.13 Other

*If results are available concerning any other vitamins not listed here, then these should be reported at the end of this section. In addition, where information is available concerning multiple vitamin deficiencies, these should be reported at the end of this section under 'other'.*

### 5.6 Minerals

- 5.6.1 Sodium
- 5.6.2 Iron
- 5.6.3 Calcium (and Vitamin D)

- 5.6.4 Selenium
- 5.6.5 Iodine
- 5.6.6 Other

*Results are likely to be available on other minerals e.g. magnesium, potassium, zinc, copper, phosphorus, manganese and chromium for certain cancers. These should be reported at the end of this section when appropriate under 'other'.*

## 5.7 Phytochemicals

- 5.7.1 Allium compounds
- 5.7.2 Isothiocyanates
- 5.7.3 Glucosinolates and indoles
- 5.7.4 Polyphenols
- 5.7.5 Phytoestrogens eg genistein
- 5.7.6 Caffeine
- 5.7.7 Other

*Where available report results relating to other phytochemicals such as saponins and coumarins. Results concerning any other bioactive compounds, which are not phytochemicals should be reported under the separate heading 'other bioactive compounds'. Eg flavonoids, isoflavonoids, glycoalkaloids, cyanogens, oligosaccharides and anthocyanins should be reported separately under this heading.*

## 5.8 Other bioactive compounds

# 6 Physical activity

## 6.1 Total physical activity (overall summary measures)

### 6.1.1 Type of activity

- 6.1.1.1 Occupational
- 6.1.1.2 Recreational
- 6.1.1.3 Household
- 6.1.1.4 Transportation

- 6.1.2 Frequency of physical activity
- 6.1.3 Intensity of physical activity
- 6.1.4 Duration of physical activity

## 6.2 Physical inactivity

## 6.3 Surrogate markers for physical activity e.g. occupation

# 7 Energy balance

## 7.1 Energy intake

### 7.1.1 Energy density of diet

## 7.2 Energy expenditure

# 8 Anthropometry

## 8.1 Markers of body composition

### 8.1.1 BMI

### 8.1.2 Other weight adjusted for height measures

### 8.1.3 Weight

### 8.1.4 Skinfold measurements

### 8.1.5 Other (e.g. DEXA, bio- impedance, etc)

### 8.1.6 Change in body composition (including weight gain)

## 8.2 Markers of distribution of fat

### 8.2.1 Waist circumference

### 8.2.2 Hips circumference

### 8.2.3 Waist to hip ratio

### 8.2.4 Skinfolds ratio

### 8.2.5 Other e.g. CT, ultrasound

## 8.3 Skeletal size

### 8.3.1 Height (and proxy measures)

### 8.3.2 Other (e.g. leg length)

## 8.4 Growth in fetal life, infancy or childhood

### 8.4.1 Birthweight,

### 8.4.2 Weight at one year

# 9 Diet-related complementary medicine



### **Annex 3. Tables of included and excluded biomarkers proposed by the SLR centre Bristol (SLR prostate cancer, cancer survivors).**

Extracted from: Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective  
Systematic Literature Review – Support Resource  
SLR Prostate Cancer (pp 1185-1186)

The reviewers of the SLR centre Bristol used two chapters (Willet: Nutritional epidemiology (Chapter 9), 1998; Margetts and Nelson: Design concepts in nutritional epidemiology (Chapter 7), 1997) to guide their decisions. If there was no info, the biomarker was excluded. If one of the chapters stated the biomarker was useful, the data on validity were checked. Biomarkers with a correlation  $>0.20$  were included. If the chapters stated that there were no good biomarkers for a nutrient or that the biomarker was valid for certain range of intake only, the biomarker was excluded. It was assumed that if biomarkers measured in plasma were valid, this would also be true for serum and vice versa.

The reviewers of the SLR centre Bristol have been more inclusive with respect to the validation required for biomarkers of important nutrients and have therefore added serum/plasma retinol, retinol binding protein, vit B6, ferritin, magnesium, erythrocyte superoxide dismutase (more details below). They have also included biomarkers where validity is not possible: this happens in the case of toxins and phytochemicals where dietary data are sparse. Various contaminants, such as cadmium, lead, PCBs in the serum are also included now although validity data are not available. The level of these chemicals in human tissues is often the only available measure of ingestion.

<b>Measured in</b>	<b>Include</b>	<b>Exclude</b>
Serum	Provit A carotenoids: Carotene, B-carotene, Alpha-carotene Nonprovit A carotenoids: Carotenoids, Lycopene, Cryptoxanthin (B-), Lutein+zeaxanthin Vit E: alpha-tocopherol, gamma tocopherol Selenium n-3 fatty acids: EPA (Eicosapentaenoic), DHA (Docosahexaenoic) Magnesium Vit A: Retinol & Retinol Binding Protein Pyridoxic acid (vit B6) Phytoestrogen: Genistein, Daidzein Chemical food contaminants Polychlorinated biphenyls (PCBs) Phytochemicals	Prealbumin Minerals: Zinc, Copper, Copper/zinc ratio, Zinc/retinol ratio Other dietary lipids: Cholesterol, Triglycerides Saturated fatty acids, Monounsaturated fatty acids, Polyunsaturated fatty acids Lipids (as nutrients), Total fat (as nutrients), Total protein
Urine	4-pyridoxic acid (vit B6) in 24-h urine	Nitrosamines Xanthurenic acid in 24-h urine Arsenic Ferritin
Saliva		Other dietary lipids: Cholesterol, Triglycerides
Erythrocyte	Linoleic acid Selenium Superoxide dismutase Cadmium	Minerals: Zinc, Copper Monounsaturated fatty acids n-3 fatty acids: EPA (Eicosapentaenoic), DHA (Docosahexaenoic) n-6 fatty acids (other than linoleic acid) Polyunsaturated fatty acids, Saturated fatty acids Glutathione peroxidase

<b>Measured in</b>	<b>Include</b>	<b>Exclude</b>
Plasma	Vit D Vit E: alpha-tocopherol, gamma tocopherol Vit C Provit A carotenoids: Carotene, Alpha-carotene, B-carotene Nonprovit A carotenoids: Lycopene, Cryptoxanthin (B-), zeaxanthin, Lutein Selenium, Selenoprotein Folate, Iron: ferritin Vit A Retinol: Retinol Binding Protein Cadmium, Cadmium/zinc ratio EPA DHA fatty acids	Alkaline phosphatase Minerals: Zinc, Copper, caeruloplasmin Other dietary lipids: Cholesterol, Triglycerides, LDL, HDL
Adipose tissue	n-3 fatty acids: EPA (Eicosapentaenoic), DHA (Docosahexaenoic) n-6 fatty acids Trans fatty acids , Polyunsaturated fatty acids, Saturated fatty acids	Unsaturated fat, Monounsaturated fatty acids n-9 fatty acids other measures of polyunsat fa: M:S ratio, M:P ratio, n3-n6 ratio
leucocyte	Vit C	Zinc
Erythrocyte membrane	n-6 fatty acids: linoleic	n-6 fatty acids (other than linoleic) n-3 fatty acids: EPA (Eicosapentaenoic), DHA (Docosahexaenoic)
Hair		Minerals: Zinc, Copper, Manganese, Iron Cadmium
Toenails or fingernails	Selenium	Cadmium, zinc

**Reasons for exclusion and inclusion of biomarkers proposed by the SLR centre Bristol.**

Extracted from: Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective

Systematic Literature Review – Support Resource

SLR Prostate Cancer (pp 1187-1189)

(Source: Willett: Nutritional epidemiology (Chapter 9), 1998; Margetts and Nelson: Design concepts in nutritional epidemiology (Chapter 7), 1997)

<b>Exposure</b>	<b>Measured in</b>	<b>Valid?</b>	<b>Reason (Willett)</b>	<b>Reason (Margetts / Nelson)</b>
Retinol	Plasma/serum	Yes	Can be measured adequately, but limited interpretability in well-nourished population (p 190).	Main biochemical marker of vit A intake is serum retinol (p 194) although in western countries dietary intake of this vitamin is only a very minor determinant of its plasma levels.
Retinol-Binding protein	Serum	Yes	Retinol levels are highly correlated to RBP(p192).	May be measure of physiologically available form. Not if certain disease processes exist (p 192).
Beta-carotene	Plasma	Yes	Yes (p 194) although blood levels much more responsive to supplemental beta-carotene than beta-carotene from food sources (p 193)	Yes (p 197)
Alpha-carotene Beta-cryptoxanthin Lutein+zeaxanthin Lycopene	Plasma	Yes	Yes (p 194)	There is some evidence for interaction between carotenoids during intestinal absorption, which may complicate relationship between intake and blood levels (p 198)
Vit E	Plasma	Yes	Yes (p 196) NB. Strong confounding with serum cholesterol and total lipid concentrations (p 196).	Plasma, red and white blood cells. Yes, if used for vit E supplements. Yes, although if used for diet, associations are only moderate (p199)

Exposure	Measured in	Valid?	Reason (Willett)	Reason (Margetts / Nelson)
Vit D: D25 (OH)D	Plasma Serum	Yes	Yes (P 198/199) NB. Seasonal variation exists, especially in elderly populations, decreasing in winter and rising during summer (p 198) Sunshine exposure is most important determinant; level is better marker of dietary intake in subjects with low sun exposure	Both can be used to measure vit D status, but the higher plasma concentration and lesser metabolic control of d25 makes this, by far, the better option (p 198).
Vit D: 1.25 (OH)2D		No	No. Influenced by calcium and phosphate levels and parathyroid hormone (p 199).	
Vit D: Alkaline phosphatase activity	Serum	No	No. Is indirect measure of vit D status and is susceptible to other disease processes (p 199)	No info
Vit C	Plasma Leukocyte Serum	Yes	Yes (p 200). Leukocyte may be preferred for long-term intake and plasma and serum reflects more recent intake (p 201)	Yes (p 209), vit C exhibits the strongest and most significant correlation between intake and biochemical indices. Known confounders are: gender, smoking
Vitamin B6	Plasma	Yes	Yes response to supplementation shows response in PLP. PLP better measure of short term rather than long term	Recent studies show that there is unlikely to be a strong correlation between dietary intake and plasma pyridoxal phosphate levels (PPL)
PLP and 4 Pyridoxic acid	Urinary	Yes	Urinary B6 may be more responsive to recent dietary intake than plasma PLP. Random samples of urine 4-pyridoxic acid correlate well with 24 hour collections	
Folacin (folate)	Serum Erythrocyte	Yes	Yes good correlation with dietary folate in both serum and erythrocytes	Used for assessing folate status Table 7.11p
Magnesium	Serum	Yes	Yes stronger correlation with supplement users than with dietary Mg	
Iron	Serum Hair/nails	No No	No, short-term variability is very high (p 208). No, remains to be determined	
Iron: Ferritin	Serum	Yes	Meat intake predicts serum ferritin level (p 208)	No marker of iron intake is satisfactory (p. 192)

Exposure	Measured in	Valid?	Reason (Willett)	Reason (Margetts / Nelson)
Copper : Superoxide dismutase	Erythrocyte	Yes	Among four men fed a copper deficient diet for 4 months, erythrocyte S.O.D declined for all 4. Copper repletion restored S.O.D levels	
Copper	Plasma/serum	No	No (p 211): large number of lifestyle factors/pathologic conditions probably alter blood copper concentrations (smoking, infections)	
Copper	Hair	No	No evidence (212) and data suggests influenced by external contamination	No. Copper-dependent enzyme superoxide dismutase in erythrocytes and copper-protein complex caeroplasmin in serum have been shown to be associated with copper intake, but these markers may be influenced by nondietary factors (p 193)
Selenium	Blood components Toenails	Yes	Yes. Erythrocyte is probably superior to serum as measure of long-term intake (p 206). Lower influence of environment in countries where wearing shoes is norm (toenails). Selenium status is reduced by smoking, also in older persons (p 207); Relationship of selenium with disease may be modified by other antioxidants (vit E and C)	Yes (p 193). Relationship between selenium intake and biomarkers is reasonably good. Urine: reasonable marker, plasma reflects intake provided that the range of variation is large. Red cell and glutathione peroxidase are markers of longer-term intakes. Hair and toenails are alternative possibilities, although contamination of hair samples with shampoo must be controlled for
Glutathione peroxidase	Plasma Serum Erythrocytes Blood	No	Is poor measure of selenium intake among persons with moderate and high exposure (p 206)	

Exposure	Measured in	Valid?	Reason (Willett)	Reason (Margetts / Nelson)
Zinc Metallothionein levels	Any	No No	No (p 212) May be marker of short-term intake (p 213)	No biochemical marker is a good indicator of zinc intake (p 192/193). This is, in general terms, also true for other trace metal nutrients such as copper, manganese, chromium, etc
Lipids: total fats	Any	No	No (p 213)	No, there are no markers of total fat intake (p 215)
Cholesterol, LDL Lipoprotein levels	Serum	No	No, but may be useful to predict dietary changes but not for dietary intake (p 215)	No, relationship dietary cholesterol and lipoprotein levels of cholesterol are complex and appears to vary across range of intake (p218)
Linoleic acid	Plasma  Adipose tissue	No  Yes	Plasma linoleic acid can discriminate between groups with relatively large differences in intake but performs less well on an individual basis (p 220) Yes (p 220)	No consistent relation between dietary linoleic acid intake and plasma linoleic acid (p 220). Across the range of fatty acids in the diet, fatty acids levels in blood and other tissue (adipose tissue) reflect the dietary levels. NB levels are not comparable across tissues
Marine omega-3 fatty acids (EPA, DHA)	Serum Plasma Adipose tissue	Yes	Yes (p 222/223), although dose-response relation remains to be determined	
Monounsatur fatty acids (oleic acid)	Plasma Adipose tissue	No No	No, plasma levels are poor predictors of oleic acid intake, but adipose tissue may weakly reflect oleic acid intake (p. 224). Validity is too low	
Polyunsatur fatty acids	Adipose tissue	Yes	Yes (p 220)	No info

<b>Exposure</b>	<b>Measured in</b>	<b>Valid?</b>	<b>Reason (Willett)</b>	<b>Reason (Margetts / Nelson)</b>
Saturated fatty acids (Palmitic acid, stearic acids)	Adipose tissue Plasma	Yes No	Yes, long term sat fatty acid intake may be reflected in adipose tissue levels (p 224) No, levels of palmitic and stearic acids in plasma do not provide a simple index of intake (p 224).	No info
Trans-fatty acids	Adipose tissue	Yes	Yes (p 225)	No info
Protein	Any	No	No (p 226)	No info
Nitrogen	Urine	Yes	Yes, but several 24-h samples are needed to provide a stable estimate of nitrogen intake (p 227) Nitrogen excretion increases with body size and exercise and decreased caloric intake	Yes (p 219) One assumes that subjects are in nitrogen Balance



### Data on validity and reliability of included biomarkers

Extracted from: Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective  
 Systematic Literature Review – Support Resource  
 SLR Prostate Cancer (pp 1187-1189)

Nutrient	Biologic tissue	Val./reproduc	Coef	Details
Retinol	Plasma	Validity	0.17	Borderline Correlation between pre-formed vit A intake and plasma retinol. However plasma retinol is a recognized marker of vit A nutritional status for undernourished populations
Beta-carotene			0.51	Correlation between plasma beta-carotene level (averaged from 2 samples taken 1 week apart) and a 7-day diet record estimate of beta-carotene in 98 non-smoking women (Willett, p 194).
			0.38	Cross-sectional correlation between dietary intake of carotene and plasma betacarotene in 902 adult females. In males (n=880): r=0.20 (Margetts, table 7.9a).
	Plasma	Reproducibility	0.45	Correlation for carotene (80% beta-carotene, 20% alpha-carotene) between two measurements taken 6 years apart (Willett, p 194).
Beta-cryptoxanthin	Plasma	Validity	0.49	Correlation between plasma beta-carotene level (averaged from 2 samples taken 1 week apart) and a 7-day diet record estimate of beta carotene in 98 non-smoking women (Willett, p 194)
Lutein+zeaxanthin	Plasma	Validity	0.31	
Lycopene	Plasma	Validity	0.50	
Alpha-carotene	Plasma	Validity	0.58	
Alpha-carotene	Plasma	Validity	0.43	Cross-sectional correlation between dietary intake of carotene and plasma alphacarotene in 902 adult females. In males (n=880): r=0.41 (Margetts, table 7.9a).
Carotenoids	Plasma	Reproducibility	≥080	Within-person variability of plasma levels over 1 week (Willett, p 194).
Vitamin E	Plasma	Validity	0.53	Lipid-adjusted alpha-tocopherol measurements and estimated intake (incl. supplements). After excluding supplement users: r=0.35 (Willett, p 196)
	Plasma	Reproducibility	0.65	Unadjusted repeated measures over a 6-year period (p 188). Adjusting for serum cholesterol reduced correlation to r=0.46 (p 188). Also r=0.65 was found over a 4-year period in 105 adults in Finland (Willett, p 196).
	Plasma	Validity	0.20	Cross-sectional correlation between dietary intake of vit E and plasma vit E in 880 adult males. In females (n=906): r=0.14 (Margetts, table 7.9a)

Nutrient	Biologic tissue	Val./reproduc	Coef	Details
Vitamin D: D25 (OH)D	Plasma	Validity	0.35	Correlation between FFQ estimate of vit D intake (including supplements) with plasma D25 (OH)D (n=139). Correlation excluding supplement users: r=0.25 (Willett, p 199)
			0.18	Cross-sectional correlation between dietary intake of nutrients and biochemical markers in UK pre-school child study in females (n=350). In males (n=365) r=0.06 (Margetts, table 7.9b).
	Serum	Validity	0.24	Correlation between estimated vit D intake from food and supplements (based on 24 h recall) and serum D25 (OH)D (n=373 healthy women). Food only: r=0.11 (Willett, p 199).
Vitamin C	Plasma	Validity	0.43	Unadjusted correlation between questionnaire-derived dietary ascorbic acid intake and plasma ascorbic acid concentration in a heterogeneous population. Diet only: r=0.38 (Table 9.1). Correlation is 0.31 for leukocyte ascorbic acid concentration. (Willett, p 200)
		Reproducibility	0.28	Repeated measures in men obtained 6 years apart (Willett, p 201)
		Validity	0.43	Cross-sectional correlation between dietary intake of nutrients and biochemical markers in UK pre-school child study in males (n=369). In females (n=354) r=0.39 (Margetts, table 7.9b).
	Serum	Validity	0.55	Correlation between food-frequency questionnaire estimate of vit C intake and serum vit C values (in smokers) in 196 men in Scotland (adjusted for total energy intake, BMI and serum cholesterol level). Non-smokers: 0.58 (Willett, p 200/201)
	Leukocyte	Validity	0.49	Correlation between one week of intake data and a single leukocyte ascorbate measurement for men. For women: r=0.36. Nutrition survey of elderly in UK (Margetts, p 211)
Vitamin B6	Plasma	Validity	0.37	Correlation between B6 and plasma pyridoxal phosphate levels in 280 healthy men =0.37 (Willett p203)
	Urinary	Validity	-	
Folacin	Serum Erythrocyte	Validity	0.56	Correlation of 0.56 in Framington Heart study 385 subjects (serum)
			0.51	Correlation in 19 elderly subjects (erythrocyte) (Willett p204)
Magnesium	Serum	Validity	0.27	Correlation between intake with supplements 0.27 in 139 men and 0.15 without supplements (Willett p211)

<b>Nutrient</b>	<b>Biologic tissue</b>	<b>Val./reproduc</b>	<b>Coef</b>	<b>Details</b>
Iron (ferritin)	Serum	Validity	0.16	Borderline 0.16 correlation with heme intake but only r=0.15 with total iron intake (Willett p 208). Included as marker of iron storage
Copper (Superoxide dismutase)	Erythrocyte	-	-	S.O.D levels reflect both depletion and repletion of Cu (Willett p 212)
Selenium	Serum	Validity	0.63	Correlation between selenium intake and serum selenium in South Dakotans (n=44)(Willett, p 186)
		Reproducibility	0.76	Average correlation between repeated measurements at four 3-month intervals in 78 adults (Willett, p 188)
	Toenails	Validity	0.59	Correlation between selenium intake and toenail selenium level in South Dakotans (n=44) (Willett, p 186)
		Reproducibility	0.48	Correlation for selenium levels in toenails collected 6 years apart from 127 US women (Willett, p 206)
	Whole blood	Validity	0.62	Correlation between selenium intake and whole blood selenium in South Dakotans (n=44) (Willett, p 186)
		Reproducibility	0.95	Average correlation between repeated measurements at four 3-month intervals in 78 adults (Willett, p 188)
Linoleic acid	Adipose tissue	Validity	0.57	Correlation between dietary linoleic acid intakes determined from 7-day weighted diet records and the relative proportion of linoleic acid in adipose tissue in Scottish men (n=164). Also correlation between linoleic acid measured in adipose tissue and calculated from FFQ in 118 Boston-area men (Willett, p 220)
Eicosapentaenoic (n-3)	Adipose tissue	Validity	0.40	Correlation with intake estimated from three 7-day weighted food records (Willett, p 223).
		Reproducibility	0.68	Correlation over 8 months in 27 men and women aged 20-29 (Willett, p 223).
	Plasma	Validity	0.23	Correlation of cholesterol ester fraction and intake in 3,570 adults (Willett, p 223)
		Reproducibility	0.38	Correlation of two measurements taken 6 years apart in study of 759 Finnish youths (Willett, p 219)

<b>Nutrient</b>	<b>Biologic tissue</b>	<b>Val./reproduc</b>	<b>Coef</b>	<b>Details</b>
Docosahexaenoic (n-3)	Adipose Tissue	Validity	0.66	Correlation with intake estimated from three 7-day weighted food records (Willett, p 223)
		Reproducibility	0.93	Correlation over 8 months in 27 men and women aged 20-29 (Willett, p 223).
	Plasma	Validity	0.42	Correlation of cholesterol ester fraction and intake in 3,570 adults (Willett, p 223)
		Reproducibility	0.38	Correlation of two measurements taken 6 years apart in study of 759 Finnish youths (Willett, p 219)
Polyunsaturated fatty acids	Adipose tissue	Validity	0.80	Correlation between % of polyunsaturated fatty acid relative to total fatty acid intake and relative % of adipose tissue polyunsaturated fatty acid (Willett, p 220)
Palmitic acid	Adipose tissue	Validity	0.27	Correlation adipose tissue measurement with a FFQ estimate among 118 men. A correlation of 0.14 was reported among women. Among 20 healthy subjects, correlations between normal intake of total saturated fatty acids and fatty acid composition of triglycerides in adipose tissue was 0.57 (Willett, p 224)
Stearic acid	Adipose tissue	Validity	0.56	Among 20 healthy subjects, correlations between normal intake of total saturated fatty acids and fatty acid composition of triglycerides in adipose tissue (Willett, p 224)
Trans fatty acids	Adipose tissue	Validity	0.40	Correlation between adipose trans and intake estimated from the average of two FFQ among 140 Boston-area women. Previous study: 115 Boston area women, correlation of 0.51 between trans intake estimated from a single FFQ and a fatty acid measurement. Among 118 Boston-area men: correlation of 0.29 between trans fatty acid measured in adipose and by FFQ (Willett, p 225)
Nitrogen	Urine	Validity	0.69	Correlation between nitrogen intakes estimated from weighted food records of 16 days and the average of six 24-h urine nitrogen levels (160 women) (Willett, p 227)
Phyto Oestrogens Genistein, daidzein	Plasma 24 hr urine	Validity	0.97 0.92	Urinary excretion (24 h) and plasma concentrations of PO were significantly related to measured dietary PO intake (r 0.97, P<0.001 and r 0.92, P<0.001 respectively). These findings validate the PO database and indicate that 24 h urinary excretion and timed plasma concentrations can be used as biomarkers of PO intake. Br J Nutr. 2004 Mar;91(3):447-57

Nutrient	Biologic tissue	Val./reproduc	Coef	Details
Enterodiol Enterolactone	Serum Urine	Validity	0.13 to 0.29	Urinary enterodiol and enterolactone and serum enterolactone were significantly correlated with dietary fiber intake ( $r = 0.13-0.29$ ) Cancer Epidemiol Biomarkers Prev. 2004 May;13(5):698-708

## **Appendix 8. List of variables in WCRF database**

### **Paper information**

- WCRF code
- PMID
- Study type
- Author
- Publication year
- Paper title
- Journal
- Volume
- Starting page
- Ending page
- Enterer ID
- Cancer site

### **Subjects information**

- Region where the study was conducted
- Country
- Ethnicity
- Nationality
- Gender
- Mean age
- Age range
- Breast cancer characteristics
- Other characteristics (notes on co-morbidity, socio-demographics)
- Recruitment procedure

### **Study information**

- Study size
- Source of cases
- Criteria for inclusion
- Criteria for inclusion
- Response rate
- Cases/Controls ratio
- Number of cases (total)
- Number of non-cases (total)

- Study design type (Factorial, Crossover, etc)
- Matching factors
- Textual note for matching
- Number of participating centres
- Is there comparability across centres?
- Power estimation done?
- Follow-up method
- Length of follow up
- Loss to follow up
- Textual notes to loss to follow up
- Type of analysis
- Was clustering accounted for?
- Number of cluster groups
- Person-years

### **Exposure assessment (for each exposure)**

- Dietary Assessment method
- Dietary Assessment method (notes)
- Biomarker sample origin
- Biomarker sample (notes)
- Biomarker average years of collection to diagnosis
- Physical Activity assessment method
- Physical Activity assessment method (notes)
- Anthropometry assessment method
- Anthropometry assessment method (notes)
- Measurement error accounted for

### **Interventions**

- Intervention type
- Number of intervention groups
- Intervention description
- Intensity/dosage
- Frequency
- Mode of delivery
- Duration
- Group size
- Randomisation
- Blinding

- Textual notes to blinding
- Study design type (Factorial, Crossover, etc)
- Length intervention
- Textual note for length intervention
- Missing participants
- Reasons for missing participants
- Other concerns about bias
- Adverse effects (notes)

### **Results (for each exposure/intervention)**

- Exposure
- Textual notes to the exposure
- Outcome type
- Outcome sub-type
- Outcome definition
- Outcome ascertainment
- Outcome (primary,secondary, ancillary)
- Time points collected
- Textual notes to outcome
- Type of result (quantiles, categories, continuous, means)
- Type of adjustment (maximally, intermediate, minimally, unadjusted)
- Is it a best model? (yes/no)
- Is it a sub-group analysis (yes/no)
- Sub-group
- Number of cases in the analysis
- Number of controls in the analysis
- Unexposed group
- Intervention group
- “Quantiles” calculated from (all population, etc)
- Adjustment factors
- Textual notes to this result
- Missing participants
- Intention-to treat analysis (notes)

### **For each “quantile”:**

- Quantile description
- Bottom range of quantile



- Top range of quintile
- Median/Mean range of quantile
- Number of cases by quantile
- Number of controls by quantile
- Event rate baseline
- Sum of cases and controls
- Type of result (RR, OR, mean difference)
- RR
- 95% CI
- p Value
- Mean cases
- Mean controls
- Mean cases SD
- Mean cases 95 % CI
- Mean p Value
- Coefficient correlation
- Coefficient correlation 95 % CI
- Slope
- Slope 95% CI
- Slope p Value

## Appendix 7. Checklist for assessment of study quality and risk of bias

### NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE

#### COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

#### Selection

##### 1) Selection of the non exposed cohort

- a) drawn from the same community as the exposed cohort \*
- b) drawn from a different source
- c) no description of the derivation of the non exposed cohort

##### 2) Ascertainment of exposure (for each exposure)

- a) secure record (eg measured by a trained person, biomarker) \*
- b) structured interview \*
- c) written self report
- d) no description

We will add the following criteria for ascertainment of exposure:

- a) Were methods of exposure ascertainment validated?\*
- b) Was there any intent to correct for measurement error of exposure (e.g. repeated measures, calibration) \*

##### 3) Demonstration that outcome of interest was not present at start of study

- a) yes \*
- b) no

#### Comparability

##### 1) Comparability of cohorts on the basis of the design or analysis

a) study controls for age, systemic adjuvant therapeutic received, tumor stage, estrogen and progesterone receptor status, menopausal status (when pre and post-menopausal women are included in the same analysis) \*

b) study controls for any additional factor \*

**Outcome** (for each outcome).

1) Assessment of outcome (we added point b because interventions like physical activity and dietary changes can't be masked)

a) independent blind assessment \*

b) independent not-blinded assessment (for exposures that can't be masked) \*

c) record linkage \*

d) self report

d) no description

2) Was follow-up long enough for outcomes to occur

a) yes (more than two years) \*

b) no

3) Adequacy of follow up of cohorts

a) complete follow up - all subjects accounted for \*

b) subjects lost to follow up unlikely to introduce bias - small number lost (less than 10 % or description provided of those lost) \*

c) follow up rate loss more than 10% and no description of those lost

d) no statement

***We will exclude from the assessment::***

1) Representativeness of the exposed cohort

a) truly representative of the average \_\_\_\_\_ (describe) in the community \*

b) somewhat representative of the average \_\_\_\_\_ in the community \*

c) selected group of users eg nurses, volunteers

d) no description of the derivation of the cohort



ASSESSMENT OF TRIAL QUALITY

(BASED IN THE COCHRANE COLLABORATION'S TOOL FOR ASSESSING RISK OF BIAS)

**1. SELECTION BIAS**

**Randomization (select one):**

- Patients randomly allocated
- Groups randomly allocated
- Not randomized  .....
- Unclear  .....

**If randomization, method of randomization:**

- Central allocation (including telephone, computer, web-based and pharmacy-controlled randomization)
- Sealed envelopes
- Drug containers of identical appearance
- Other (specify)  .....
- Not clear (free text)  .....

**2. PERFORMANCE AND DETECTION BIAS**

---

	Yes	No	Unclear	Intervention can't be masked
Were participants blind to the intervention assigned	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Were the clinicians/investigators blind to which intervention was being provided?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Were the assessors of outcome measures blind to the intervention provided?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Were the statisticians blind to the intervention provided?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

---

### 3. LOSSES TO FOLLOW UP

For each outcome:

	Number	%	Not reported
<hr/>			
How many subjects were lost to outcome follow up?			
<hr/>			
	Yes	No	Unclear
Are subject lost to follow up enumerated with reasons for loss?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Are subject loss to follow up included in the denominator for analyses of outcomes?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Were outcome missing data imputed or outcome measurements adjusted for losses to follow up?	<input type="checkbox"/> Yes	<input type="checkbox"/> No	<input type="checkbox"/> Unclear
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 4. SELECTIVE REPORTING

<b>Selective outcome reporting</b>	Yes	No	Unclear
Are reports of the study free of suggestion of selective outcome reporting?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Was the study apparently free of other problems that could put it at a high risk of bias?			

## Appendix 8: Procedures to calculate missing information for meta-analysis<sup>14</sup>

Type of data	Problem	Assumptions
Dose-response data	Serving size is not quantified or ranges are missing, but group descriptions are given	Use serving size recommended in SLR <sup>1</sup> for consistency in the analyses (Appendix 4)
	Standard error missing	The p value (either exact or the upper bound) is used to estimate the standard error
Quantile-based data	Numbers of controls (or the denominator in cohort studies) are missing	Group sizes are assumed to be approximately equal
	Odds ratio is missing	Unadjusted odds ratios are calculated by using numbers of cases and controls in each group
	Confidence interval is missing	Standard error and hence confidence interval were calculated from raw numbers (although doing so may result in a somewhat smaller standard error than would be obtained in an adjusted analysis)
	Group mean or median are missing	The mid-point of closed-ended categories will be assigned as exposure to the group. The median exposure for open ended-categories will be estimated by using the method of Chene and Thompson <sup>11</sup> assuming a normal or lognormal distribution. However, if the number of groups is too small to calculate a distribution, the midpoint will be assigned to the lowest category. The upper bound plus the mid-range of the precedent category will be assigned to the highest category.
Category data	Numbers of cases and controls (or the denominator in cohort studies) is missing	These numbers may be inferred based on numbers of cases and the reported odds ratio (proportions will be correct unless adjustment for confounding factors considerably alter the crude odds ratios)

