Protocol *Version 2*

Continuous update of the epidemiological evidence on food, nutrition, physical activity and the risk of gastric cancer.

Prepared by: CUP team, Imperial College London, October 2012
Revised in February 2013

Introduction

The World Cancer Research Fund/ American Institute for Cancer Research: (WCRF/AICR) has been a global leader in elucidating the relationship between food, nutrition, physical activity and cancer. The first and second expert reports (1;2) represent the most extensive analyses of the existing science on the subject to date. The second expert report was informed by a process of systematic literature reviews (SLRs) all of the evidence published. Seventeen SLRs were carried out in different centres and the information collected was stored in one database for each of the cancer sites that were reviewed.

The second report features eight general and two special recommendations based on solid evidence which, when followed, will be expected to reduce the incidence of cancer. A recent study in a large European cohort study showed that people with lifestyle in agreement with the WCRF/AICR recommendations experienced a decreased risk of cancer after an average follow-up time of ten years. The main risk reductions were observed for cancers of the colon and rectum, and stomach cancer, and significant associations were observed for cancers of the breast, endometrium, lung, kidney, upper aerodigestive tract, liver, and oesophagus but not for prostate, ovarian, pancreatic, and bladder cancers (3).

To keep the evidence current and updated into the future, WCRF/AICR is undertaking the Continuous Update Project (CUP) in collaboration with Imperial College London (ICL). The CUP [http://www.wcrf.org/cancer_research/cup/index.php] is an on going review of nutrition and cancer research on food, nutrition, physical activity and body fatness, and cancer risk that captures and reviews the evidence as it accumulates. The project ensures that the evidence on which the WCRF/AICR recommendations are based continues to be the most-up-to-date and comprehensive available.

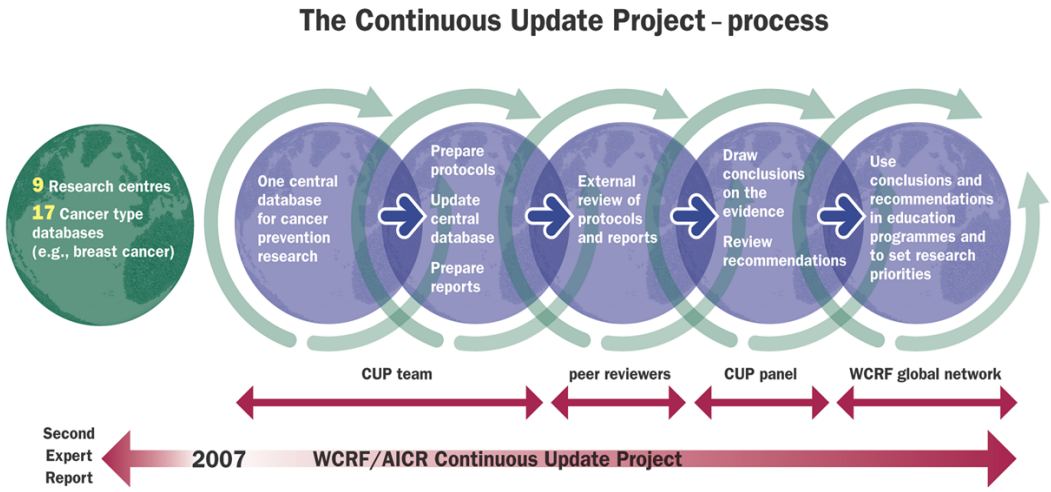
The CUP builds on the foundations of the second expert report to ensure a consistent approach to reviewing the evidence and it follows the methods developed specifically for the Second Expert Report. The methods are detailed in the SLR Specification Manual (4).

The CUP is conducted by a team at ICL, where a central database has been created by merging the cancer-specific databases generated during the SLR's in the participating centres. A key step of the CUP is to update the central database with evidence published since the Second Expert Report. The meta-analyses conducted for the

Second Expert Report will be updated by adding the new evidence identified during the CUP to the evidence collected in the 2007 SLRs.

WCRF/AICR has convened a panel of experts for the CUP consisting of leading scientists in the field of diet, physical activity, obesity and cancer, who will consider the evidence produced by the systematic literature reviews and meta-analyses, and draw conclusions before making recommendations. The entire CUP process will provide an impartial analysis and interpretation of the data as a basis for reviewing and where necessary revising the 2007 WCRF/AICR's cancer prevention recommendations (**Figure 1**).

Figure 1. The Continuous Update Process



The evidence of the different cancers is being updated progressively in a rolling programme. The CUP started in 2007 and breast cancer was the first cancer to be updated, followed by prostate cancer, colorectal cancer and other cancer sites. When a cancer site is included in the review, the CUP team at ICL keeps updating the database for that cancer and all the other cancers already included in the CUP (**Figure 2**). Currently, the central database is up-to-date for cancers of the breast, prostate, colon and rectum, pancreas, ovary, endometrium, bladder, kidney, gallbladder and liver.

Periodically, the CUP team at ICL prepares reports on the relationship of foods, nutrition, physical activity and body weight by request of the CUP Panel and the Secretariat of the project. The CUP team at ICL has completed updated reports on cancers of the breast, colon and rectum, and pancreas.

The protocols and reports of the CUP are available at http://www.dietandcancerreport.org/cancer_resource_center/continuous_update_project.php).

The present document is the protocol for the continuous update of the epidemiological evidence on food, nutrition, physical activity and the risk of stomach cancer. The peer-reviewed protocol will represent the agreed plan for the continuous update.

Should departure from the agreed plan be considered necessary at a later stage, the CUP Expert Panel must agree this and the reasons documented.

Figure 2. The Continuous Update Project- rolling programme

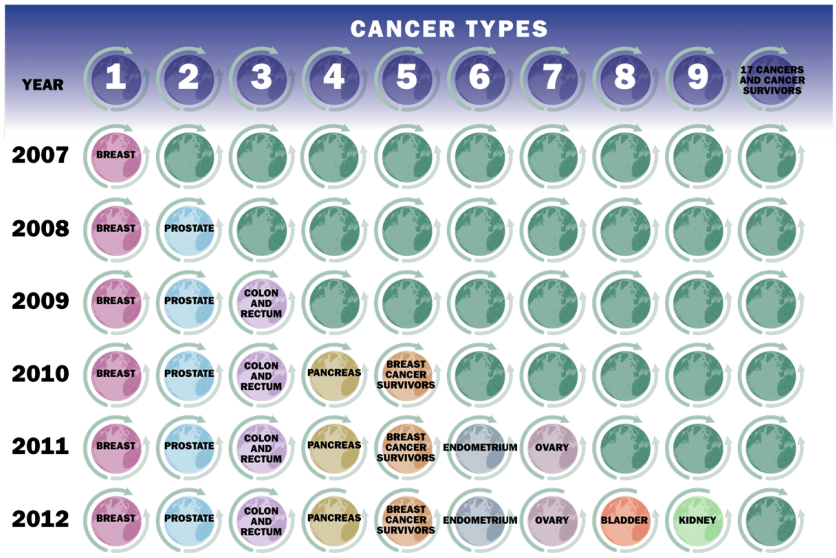
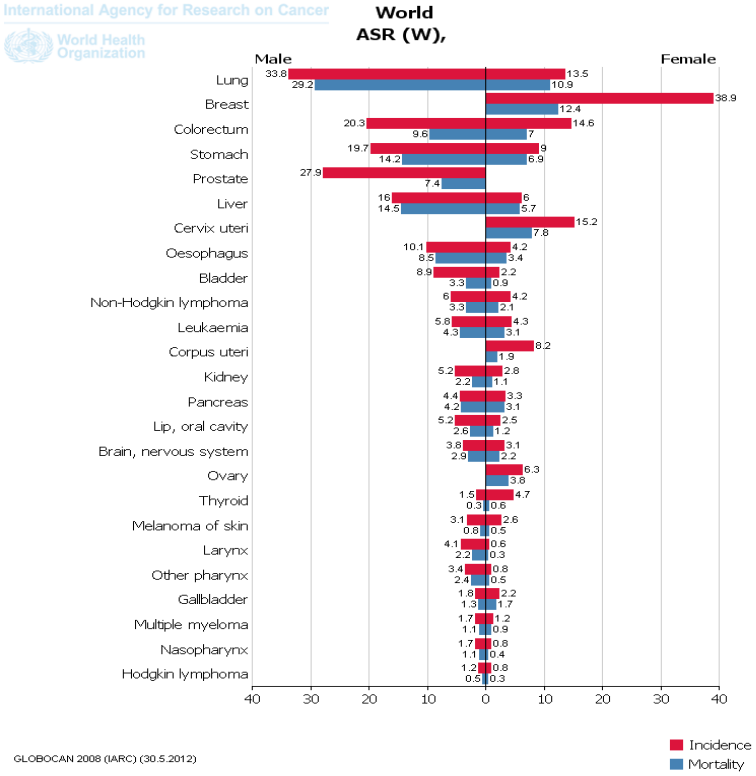


Figure 3. Estimated age (world)-standardized incidence and mortality rates of most frequent cancers (per 100 000) by sex. World. 2008



- **Gastric cancer: Epidemiology and clinical aspects**

Gastric cancers, also called stomach cancer, are cancer that forms in tissues lining the stomach. Gastric cancer is the fourth most common malignancy and the second leading cause of death due to cancer worldwide. In 2008, more than 990,000 cases were recorded (7.8% of new cancers) with 738,000 deaths (5) **(Figure 3).** Gastric cancer has a poor prognosis as it is usually diagnosed at an advanced stage. In many populations, age-standardized incidence rates are about twice as high in men as in women.

Gastric cancer has two main anatomical localizations in the stomach: tumours arising in the cardia – upper portion of the stomach that adjoins the opening of the oesophagus into the stomach- and those from distal stomach (non-cardia).

The vast majority of gastric malignancies are adenocarcinomas, which are commonly divided into intestinal type and diffuse (undifferentiated) type carcinomas (6). Most gastric carcinomas are of the intestinal type. Both histologic types are strongly associated with *H. pylori* infection (7).

Premalignant gastric lesions are risk factors for the development of intestinal-type gastric adenocarcinomas. A multistep sequence of the precursor lesions generally precedes these tumours, in a cascade in which *H. pylori* causes chronic inflammation of the gastric mucosa, followed by a slowly progression through the premalignant stages of atrophic gastritis, intestinal metaplasia and dysplasia to gastric adenocarcinoma. The risk for progression of *H. pylori*-induced gastritis toward premalignant lesions and gastric cancer depends on the duration, distribution, and severity of chronic active *H. pylori* gastritis. [reviewed by de Vries and Kuipers (8)].

The highest incidence rates of gastric cancer are observed in some countries from Eastern Asia, South America and Eastern Europe **(Figure 4).** The highest age-standardised incidence rates for both sex combined are in the Republic of Korea (41.4 per 100, 000), Mongolia (34.0 per 100,000), Japan (31.1 per 100,000) and China (29.8 per 100,000) (5).

The incidence of gastric cancer has declined over the past 50 years in most Western countries. However, while the incidence of non-cardia gastric cancer has declined in most countries, the rates of cardia cancer has remained stable, or rose in several European countries, Japan and North America (9)

- **Gastric cancer: Risk factors**

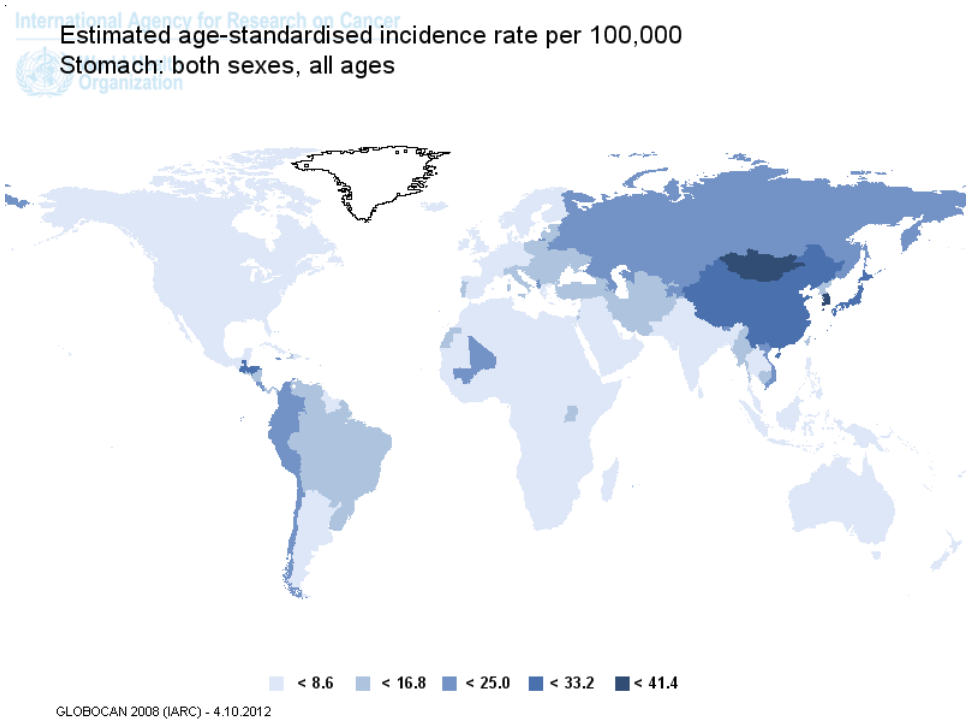
H. pylori has been classified as carcinogenic to humans Group 1 by the International Agency for Research on Cancer (10) and it is considered the single most common cause of gastric cancer. *H.pylori* infection is strongly associated with cancers located in the distal stomach (non-cardia), whereas no association has been observed for tumours located in the cardia (7)

Tobacco smoking is considered a risk factor of gastric cancer. Between 11 and 18% of gastric cancer cases are estimated to be attributable to smoking (11).

There is evidence showing that fruits and vegetables probably decrease the risk of stomach cancer and that high salt intake probably increases it (1;12;13). Other

nutritional factors that have been found related to gastric cancer are processed meat intake and grilled, broiled and barbecued meats but the evidence is not convincing (1).

Figure 4. Estimated age (world)-standardized incidence and mortality rate of stomach cancer per 100 000. World. 2008



- **Judgement of the WCRF-AICR second export report on stomach cancer**

In the judgement of the Panel of the WCRF-AICR second expert report, non-starchy vegetables, including specifically allium vegetables, as well as fruits probably protect against stomach cancer. Salt and salt-preserved foods are probably causes of this cancer. There was limited evidence suggesting that pulses (legumes), including soya and soya products, and foods containing selenium protect against stomach cancer. The evidence suggesting that chilli, processed meat, smoked foods, and grilled (broiled) and barbecued (charbroiled) animal foods are causes of stomach cancer was judged as limited (**Figure 5**).

Figure 5. Summary of judgements of the 2007 Second Expert Report on stomach cancer (1)

FOOD, NUTRITION, PHYSICAL ACTIVITY AND CANCER OF THE STOMACH		
In the judgment of the Panel, the factors listed below modify the risk of cancer of the stomach. Judgement are graded according to the strength of the evidence		
	DECREASE RISK	INCREASE RISK
CONVINCING		
PROBABLE	Non-starchy vegetables[1] Allium vegetables[1] Fruits[1]	Salt[2] Salted and salty foods
POSSIBLY	Pulses (legumes)[3] Foods containing selenium4	Chilli[1] Processed meat[5] Smoked foods[6] Grilled (broiled) or barbecued (charbroiled) animal foods[6]
LIMITED- SUGGESTIVE	Cereals (grains) and their products; dietary fibre; potatoes; starchy roots, tubers, and plantains; nuts and seeds; herbs, spices, and condiments; meat (unprocessed); poultry; eggs; milk and dairy products; fats and oils; total fat; fatty acid composition; cholesterol; sugars; sugar (sucrose); fruit juices; coffee; tea; alcohol; dietary nitrate and nitrite, *N*-nitrosodimethylamine; drying or dried food; protein; thiamin; riboflavin; vitamin C; vitamin D; multivitamin/mineral supplements; calcium; iron; selenium supplements; carotenoids; culturally defined diets; meal frequency; eating speed; body fatness; energy intake	
SUBSTANTIAL EFFECT ON RISK UNLIKELY	None identified	

[1] Judgements on vegetables and fruits do not include those preserved by salting and/or pickling.
[2] 'Salt' here means total salt consumption, from processed foods, including salty and salted foods, and also salt added in cooking and at the table.
[3] Including soya and soya products.
[4] Includes both foods naturally containing the constituent and foods, which have the constituent added (see chapter 3.5.3).
[5] The term 'processed meat' refers to meats preserved by smoking, curing, or salting, or addition of chemical preservatives.
[6] The evidence is mostly from meats preserved or cooked in these ways.

1. Research question

The research topic is:

The associations between food, nutrition and physical activity and the risk of stomach cancer.

The main objective is:

To summarize the evidence from prospective studies and randomised controlled trials on the association between foods, nutrients, vitamin, minerals, physical activity, overweight and obesity with the risk of stomach cancers in men and women.

2. Review team

Name	Current position at IC	Role within team
Teresa Norat	Principal Research Fellow	Principal investigator
Doris Chan	Research Assistant	Supervisor of data extraction. Data analyst, report preparation
Ana Rita Vieira	Research Assistant	Data analyst, report preparation
Deborah Navarro	Research Assistant	Systematic search, article selection, data extraction
Leila Abar	Research Assistant	Systematic search, article selection, data extraction
Snieguole Vingeliene	Research Assistant	Systematic search, article selection, data extraction

Review coordinator, WCRF: Rachel Thomson
Statistical advisor: Darren Greenwood, senior Research Lecturer, University of Leeds

All the reviewers have been trained in the procedures for literature search, data selection and data extraction. The reviewers that will conduct the data analyses are trained in statistical methods for meta-analyses and have conducted several systematic reviews in the CUP that have been published in peer reviewed journals (14-25).

3. Timeline

The SLR's for the Second Expert Report ended in December 30th 2005. All the data from relevant articles published up to this date was extracted by the SLR centre for the Second Expert Report. The continuous update will search and extract data of the articles from prospective studies and randomised controlled trials published from January 1st 2006. The reviewers will verify that there are not duplicities in the database using a module for article search that has been implemented in the interface for data entry.

List of tasks and deadlines for the continuous update on stomach cancer:

Task	Deadline
Start Medline search of relevant articles published from January 1st 2006	1st December 2012
Review title and abstracts of articles identified in initial electronic search (initial search will include articles added in Medline up to 31st December 2012). Select papers for complete review	15th January 2013
Review relevant papers. Select papers for data extraction	30th January 2013
Start data extraction	1st February 2013
Hand search of references	Monthly
Continuous Medline search of relevant articles included in Medline after 31st December 2012	Monthly
Continuous selection of relevant papers based on title, abstract or complete review.	Monthly
Start quantitative analysis of articles published up to 30th March 2014*	1st May 2014
Start report writing	1st September 2014
Send report for review to CUP secretariat	30th October 2014
Review and modify report according to reviewer's comments	31th January 2015
Send reviewed report to CUP secretariat	31th January 2015
Transfer Endnote files to CUP Secretariat	31th January 2015

*For the intermediate report to the CUP Panel, end date of search will be March 30th 2014

4. Search strategy

4.1. Search database

The search aims to identify all types of evidence relevant to the research question. The Medline database (includes coverage from 70 countries) will be searched using PubMed as platform. The rationale for searching only in Medline is that the results of the SLR's for the Second Expert Report indicated that searching in databases other than Medline was not cost effective (26). Central and ClinialTrials.gov will be searched for evidence of trials relevant to this review.

4.2. Hand searching for cited references

The review team will also hand search the references of reviews and meta-analyses identified during the search.

4.3 Search strategy for PubMed

The CUP review team will use the search strategy established in the SLR Guidelines for the WCRF-AICR Second Expert Report(4). The full search strategy is in **Annex 1.**

The search will be conducted in three steps:

1) Searching for studies relating to food, nutrition and physical activity
2) Searching for all studies relating to stomach cancer:
3) Searching for all studies relating food, nutrition and physical activity, and stomach cancer

The detailed search strategy is in Appendix 1.

5. Study selection criteria for the update

5.1 Inclusion criteria

The articles to be included in the review:

- Have to present results on an exposure/intervention relevant to the review. The detailed list of exposures/interventions is in **Annex** 2.
- Must have as outcome of interest incidence or mortality of gastric (stomach) cancer, cardia or noncardia gastric cancers
- Have to present results from an epidemiologic study in men and women of one of the following types†:
  - Randomized controlled trial
  - Group randomized controlled trial (Community trial)
  - Prospective cohort study
  - Nested case-control study
  - Case-cohort study
  - Historical cohort study
- Have any publication date¶

† The references of case-control studies will be stored in a Reference Manager database, but the study results will not be extracted in the central database (see Section 6).

¶ The review team will search and extract data from articles included in Medline from January 1st 2006, closure date of the database for the Second Expert Report. Any articles missing in the 2007 SLR that may be identified by screening articles references will be included independently of publication date.

5.2 Exclusion criteria

- Studies with cases of different anatomical localisations in addition to gastric cancer. For instance, gastrointestinal cancer, gastro-oesophageal cancers, etc.

- Cohort studies in which the only measure of the relationship between the relevant exposure and outcome is the mean difference of exposure (this is because the difference is not adjusted for main confounders).
- Articles in foreign language if cannot be translated (excluding articles in Chinese, French, Italian, Spanish, Portuguese and Iranian because at members in the review team can read these languages).

6. Article selection

All references obtained with the search in PubMed will be imported in a Reference Manager Database using the filter Medline.

Additionally, customized fields will be implemented in the RefMan database (see Section 6.1).

The article selection will follow three steps:

1. An electronic search will be undertaken within Reference Manager to facilitate the identification of irrelevant records. The titles and abstracts of the articles identified by the search in Reference Manager will be the first assessed for inclusion/exclusion. This will be achieved by applying a list of terms developed and tested during the preparation of the WCRF-AICR Second Expert Report:

List of terms for use within Reference Manager Database

Radiotherapy
Chemotherapy
Cisplatinum
Docetaxel
Cell
Inhibitor
Novel
Model
Receptor
Antibody
Transgenic
Mice
Hamster
Rat
Dog
Cat
In vitro

2. In a second step, two reviewers will assess the titles and abstracts of the remaining articles. The relevance of articles in language other than English will be assessed by inspection of the title and if available in English, the abstract. If the same study is published in English and in another language, only the article in English will be kept.

3. Full papers will then be obtained for all papers for which eligibility could not be assessed by reading the title and abstract and two reviewers will then assess these papers.
4. Disagreements between the reviewers will be solved by discussion with the principal investigator.
5. If a paper reports outcomes for more than one cancer site, the reviewer will extract the data for the other cancer sites in the database, using the WCRF code of the cancers in question

6.1 Reference Manager Files

Five customized fields will be created in the reference manager database. They will be used to indicate if the study was selected upon reading of title, abstract, or entire article, the study design of included articles, the status of data extraction of the included article, the WCRF code assigned and for excluded articles, the reason for exclusion (**Table 1**)

Table 1. User-defined fields to be created in Reference Manager during article selection and data extraction.

Field	**Use**	**Terms used**	**Notes**
User Def 1	Indicate if article is relevant to the CUP review	Excludedabti; Included; excluded;	Excludedabti means excluded basing on abstract and title of the article. Without "abti" means full text is reviewed.
User Def 2	If excluded, reasons	No associations of interest; No original data/duplicates; Commentary; Foreign article in [*language*] Not adequate study design Pooled studies/meta-analyses	No associations of interest include situations such as "out of the research topic", "no measure of relationship", "no specific outcome"
User Def 3	Study design	Randomized controlled trial (RCT) Prospective cohort study Retrospective cohort study Nested case-control study Case cohort study Population-based case-control study Hospital-based case-control study Case-control study- other	The CUP only extract data from RCT, cohort/cohort based studies. Case-control studies are identified but the data is not extracted to the database.

		type of controls or control type unclear	
User Def 4	WCRF code of the article	This is done during the data extraction	WCRF codes are assigned automatically in the application when performing extraction.
User Def 5	Other notes, name of study	Indicate if includes more than one anatomical localization e.g. stomach and esophagus, gastro-oesophageal cancer, gastrointestinal cancers	

7. Data extraction

The IC team will update the WCRF-AICR central database using the interface created at Imperial College for this purpose **(Figure 6)**.

Data extracted will include study design, characteristics of study population, mean age, distribution by sex, country, recruitment year, methods of exposure assessment, definition of exposure, definition of outcome, method of outcome assessment, study size, length of follow up, lost to follow-up, analytical methods and whether methods for correction of measurement error were used.

The ranges, means or median values for each level of the exposure categories will be extracted as reported in the paper.

For each result, the reviewer will extract the covariates included in the analytical model and the matching variables. The reviewer will extract the information provided about *H.pylori* infection in the population even if this was not used as covariate in the main analysis. Measures of association, number of cases and number of comparison individuals or person years for each category of exposure will be extracted for each model used in the analyses. Stratified and subgroup analyses, and results of interaction analyses will also be extracted.

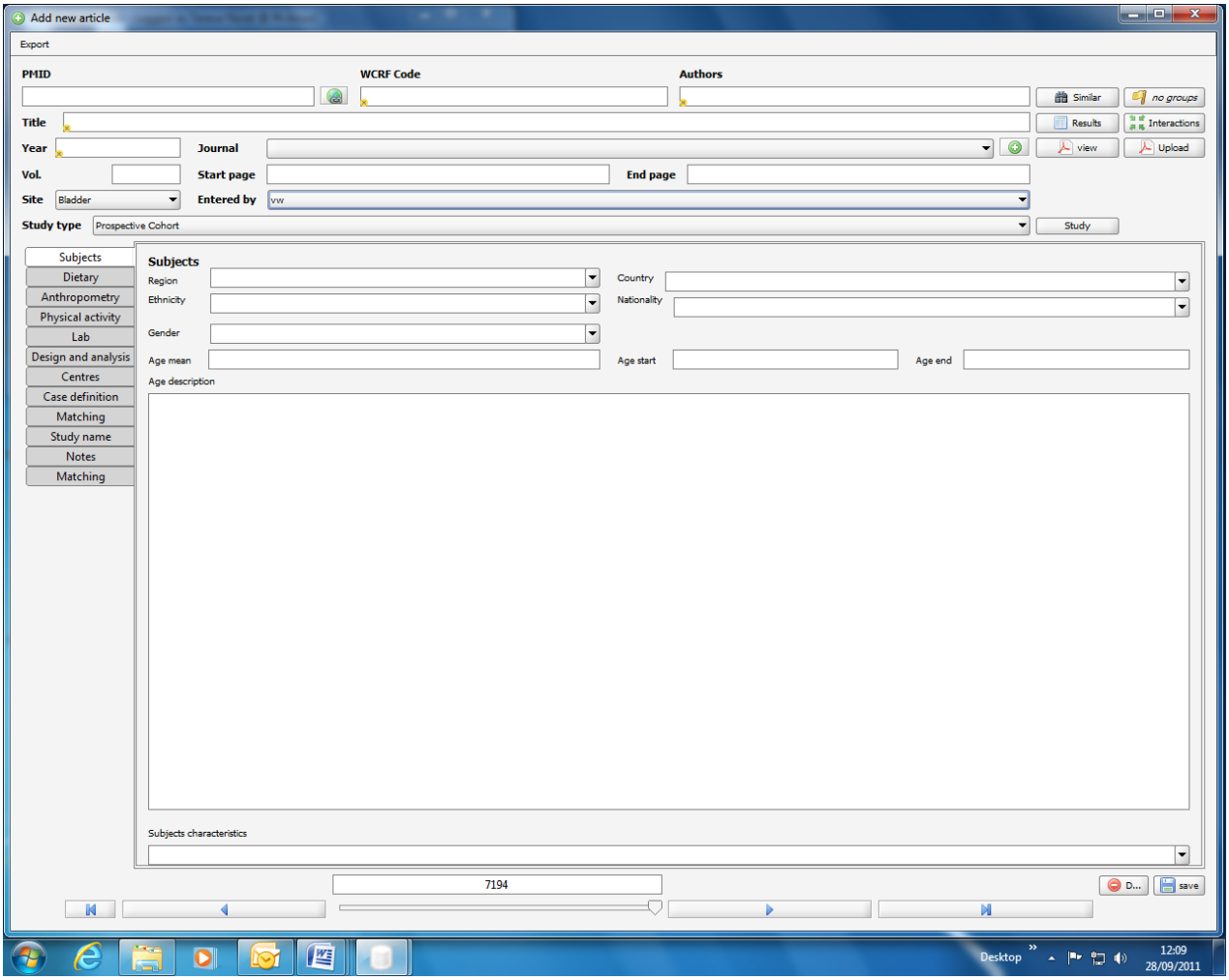
When indicated, the reviewer should also extract for each result:

- Anatomical localisation within the stomach (cardia, non-cardia)

- Histological type (adenocarcinoma, intestinal, diffuse)

- If for a subgroup or stratified analysis, the description of the subgroup or stratum

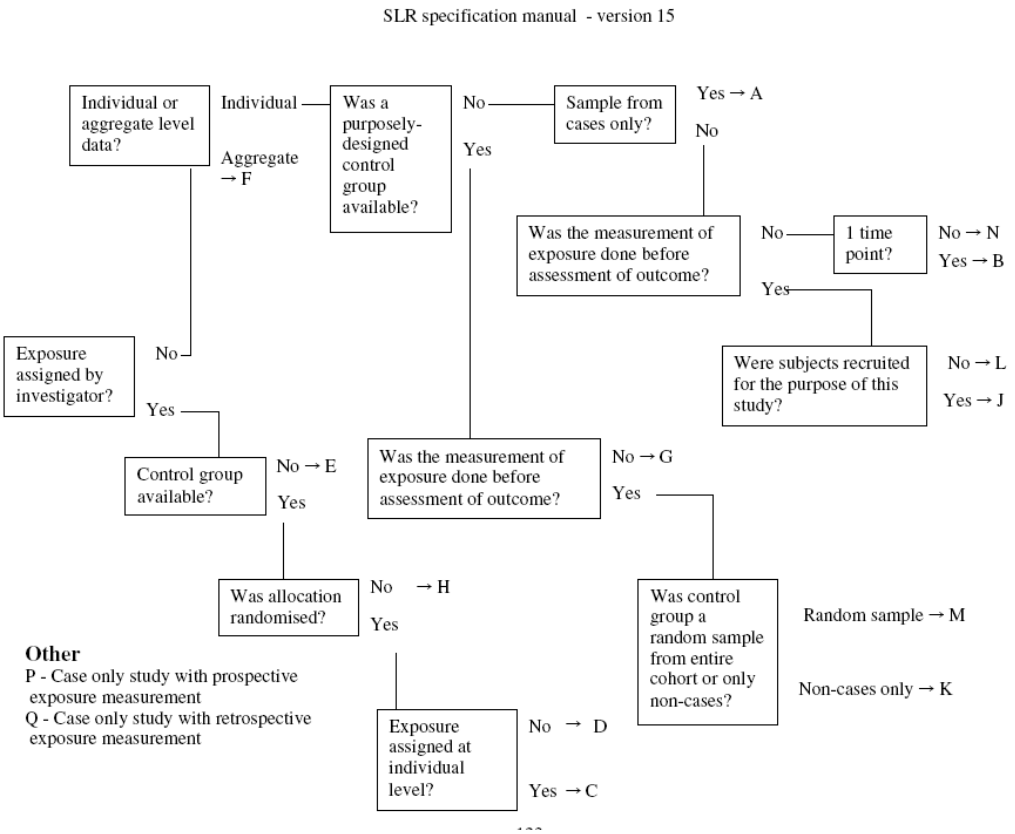
Figure 6. Example of screen for data entry. CUP



7.1 Allocation of study design

The study design algorithm devised for use of the SLR centres for the Second Expert Report will be used to allocate study designs to papers. In some cases, it will be appropriate to assign more than one design to a particular paper (e.g. analyses in the entire cohort and nested case-control). The algorithm is in **Figure 7**.

Figure 7. Study design algorithm (From: SLR specification manual)



Key to study design algorithm

Study design A Case-study / case series
Study design B Cross-sectional study
Study design C Randomised controlled trial
Study design D Group randomized control trial
Study design E Uncontrolled trial
Study design F Ecologic study
Study design G Case-control study
Study design H Non-randomized control trial
Study design J Prospective cohort study
Study design K Nested case-control study
Study design L Historical cohort study
Study design M Case-cohort study
Study design N Time series with multiple measurements

Other (see definitions in Appendix K)

Study design P Case only study with prospective exposure measurement
Study design Q Case only study with retrospective exposure measurement

7.2 Study identifier

The CUP team will use the same labelling of articles used in the SLR process for the Second Expert Report: the unique identifier for an article will be constructed using a 3-letter code to represent the cancer site: STM (stomach cancer), followed by a 5-digit number that will be allocated in sequence automatically by the interface during data extraction.

7.3 Codification of exposures/interventions.

Exposures/interventions will be codified as in the Second Expert Report for consistency. An abbreviated list of codes is in Annex 2. Additional codes for sub-exposures have been added and are programmed in the database to facilitate and standardise the data entry.

The exposures are coded by main headings and sub-headings. Wherever possible, the reviewer will use sub-heading codes. The reviewer should also extract the details of the exposure definition in the free text box in the data entry screen.

The headings for codification of the exposure groups are:

1. Patterns of diet, includes regionally defined diets, socio-economically defined diets, culturally defined diets, individual level dietary patterns, other dietary patterns, breastfeeding and other issues
2. Foods, including starchy foods; fruit and (non-starchy) vegetables; pulses (legumes); nuts and seeds; meat, poultry, fish and eggs; fats, oils and sugars; milk and dairy products; and herbs, spices, and condiments, and composite foods.
3. Beverages, including total fluid intake, water, milk, soft drinks, fruit juices, hot drinks and alcoholic drinks.
4. Food production including traditional methods and chemical contaminants, food preservation, processing and preparation.
5. Dietary constituents, including carbohydrate, lipids, protein, alcohol, vitamins, minerals, phytochemicals, nutrient supplements and other bioactive compounds
6. Physical activity, including total physical activity, physical inactivity and surrogate markers for physical activity.
7. Energy balance, including energy intake, energy density and energy expenditure.
8. Anthropometry, including markers of body composition, markers of body fat distribution, height and other skeletal measures, and growth in fetal life, infancy or childhood.

7.3.1 Codification of biomarkers of exposure

Biomarkers of exposure will be included under the heading and with the code of the corresponding exposure.

During the SLR for the Second Expert Report, some review centres opted for including in the review only biomarkers for which there was strong evidence on reliability or validity whereas other centres opted for including results on all the biomarkers retrieved in the search, independently of their validity. For the evaluation of the evidence, the Panel of Experts took in consideration the validity of the reported biomarkers.

The CUP will conduct meta-analysis for the biomarkers for which the evidence on validity and reliability was considered strong for the purpose of the Second Expert Report (full list in Annex 3). However, since the identification and validation of biomarkers is an area of research in nutritional epidemiology (27), the CUP team will extract the data for all biomarkers of intake reported in the studies, independently of whether validity and reliability had been or not fully documented.

7. 4 Extraction and labelling of study results

The reviewer will extract the associations (RR estimates and confidence intervals) with the relevant exposures from all the statistical models shown in the paper, including subgroup, stratified analyses and sensitivity analyses. These results can be presented in the paper in tables, in the text or as supplemental information.

The reviewer should label the results as unadjusted, intermediately adjusted, most adjusted model, depending of the models that are shown in the paper:

- The results for an exposure obtained with univariate models will be labelled "unadjusted".
- The results for an exposure obtained with a multivariable model including only as covariates age, sex, and in dietary analyses energy intake, will be labelled "less adjusted".
- The results for an exposure obtained with the model including the higher number of covariables in the article will be labelled "most adjusted".
- The results obtained using any multivariable model that is not the less or the most adjusted model, will be labelled "intermediately" adjusted.

In addition, the reviewer will indicate the "best model "for use in meta-analyses.

The "best" model will be the most adjusted model in the article that is a not a "mechanistic" model, which is a model that include variables likely to be in the causal pathway (e.g. milk intake as main exposure in a model adjusted for dietary calcium). When such models are reported, the "intermediately" adjusted result with the highest number of covariates will be indicated as "best model" (e.g., the most adjusted model for milk that does not include calcium).

Sometimes, potential risk factors are not kept in the final model because their inclusion in the model does not substantially modify the risk estimates. If this is indicated in the article text, this model should be considered the "best model".

In addition to adjustment, other subsidiary criteria to consider for identifying the 'best model' for meta-analysis are the completeness of the data (e.g. where number of cases is provided over where missing).

8. Quality control of the article selection and data extraction.

The article selection and the data extracted will be checked by a second reviewer at ICL. If there are discrepancies between the reviewers, the PI will decide and if there is still any doubt about the relevance of a study, the CUP Secretariat will be consulted.

9. Data analysis

9.1 Dose-response meta-analysis

Forest plots showing the study specific results for the highest versus lowest comparison exposure levels will be presented, but a meta-analytical estimate for the highest versus lowest comparison will not be calculated, to avoid pooling different exposure levels. Such as in the Second Expert Report, only linear dose-response meta-analysis will be conducted. This will allow expressing the results of each study in the same increment unit for a given exposure. In addition, non-linear dose-response meta-analyses will be conducted as exploratory analysis. In all forest plots, the studies will be ordered by publication year.

The analyses will be conducted separately for 1) cardia gastric cancer, 2) non-cardia gastric cancer and 3) studies that report on "stomach cancer" or "gastric cancer" without specifying the cancer site. Studies on cases with cancers from combined anatomical localisations will not be included (for instance, gastro-oesophageal cancers). Studies with incidence as outcome will be analysed separately from those with mortality as outcome.

Separate analyses by gender and for both gender combined will be conducted. For the analysis on both gender combined, the results for men and women in the study will be pooled first using fixed effect models and then included in the meta-analysis of "Both genders". This is essentially equivalent to including the estimate for each gender and will provide a better estimate of heterogeneity across studies.

When enough number of studies are identified, separate meta-analyses will be conducted for the subgroups reported in the papers, such as in smokers and non-smokers, with antecedents of H Pylori infection or not, and others.

Where results from two or more cohort studies are reported in the same paper, the results of each cohort will be included separately if they are provided and the pooled result will not be included. The purpose is to maintain the independence of observations included and to look at heterogeneity across study results.

The statistical methods are described in section 9.5

9.2 Selection of exposures for a dose-response meta-analysis

A dose-response meta-analysis will be conducted when at least two new reports of trials or of two cohort studies are identified during the CUP. This refers to studies providing enough information to conduct dose-response meta-analysis. The minimum number of two studies was not derived statistically but it is a number of studies that can be reasonable expected to have been published after the Second Expert Report.

The meta-analysis will include studies identified during the SLR and studies identified during the CUP. Special care will be taken to avoid including more than once the same study. Where a particular study has published more than one paper on the same exposure, the analysis using the larger number of cases will be selected but if the most recent does not provide enough information for the

dose-response meta-analysis, the publication with the required information will be selected.

If the results of the same study are not consistent across time and the most recent publication of a study cannot be included in the meta-analysis, the CUP team will conduct influence analysis of this study.

9.3 Selection of results data for meta-analyses

The results based on "best" adjusted models (full multivariable model in the articles) will be used in the dose-response meta-analyses.

When the relative risk estimate per unit of increase is reported in an article, this will be used in the CUP dose-response meta-analysis.

If the results are presented only in categorical variables (quantiles or pre-defined categories), the slope of the dose-response relationship will be derived from the categorical data.

The data required to derive the dose-response slope in each study are:

1. number of individuals with the disease for each exposure category
2. person-years -or number of individuals without the disease in nested case-control analyses- for each exposure category
3. exact cut-offs of exposure categories, or mean or median of each category.

9.4 Derivation of data required for meta-analyses.

The information provided in the articles is often incomplete and this may result in exclusions of results from meta-analyses. For instance, only 64% of the results of cohort studies on stomach and prostate cancer provided enough data to be included in dose-response meta-analysis in the SLR for the Second Expert Report. There is also empirical evidence that studies that showed evidence of an association were more likely to be usable in dose-response meta-analysis than results that did not show any evidence (28).

Failure to include all available evidence 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. To address the data incompleteness, missing data will be derived when possible during the phase of statistical analyses using other information provided in the paper (**Table 2**).

A number of approaches will be taken to derive the number of controls (or person-years) and mean exposure value for each exposure category from the available data where possible (28). When intake was expressed in "times" or "servings of intake", we will convert it into grams (g) using standard portion sizes used in the WCRF/AICR report (4) .

Means or medians of the intake categories will be assigned as "dose" when reported in the articles; if not reported, midpoints will be assigned to the relative risk of the corresponding category. If the upper boundary for the highest category was not reported, we will assume that the boundary had the same amplitude as the nearest category. For studies reporting intakes in grams/1000 kcal/day, the intake in

grams/day will be estimated using the average energy intake reported in the article. The approaches are summarized in **Table 2**.

9.5 Statistical Methods

For the linear dose-response meta-analyses, we will pool the slopes of the dose-response relationships reported in the studies. When only relative risk estimates for categorical data are reported in the paper, we will derive the slope of the "dose"-response association from the categorical data using generalized least-squares for trend estimation (29). This method accounts for the correlation between relative risks estimates with respect to the same reference category (30). The dose-response model is forcing the fitted line to go through the origin (logRR=0, dose=0). Therefore, whenever the assigned dose corresponding to the reference group (RR=1) is different from zero, all the assigned doses will be rescaled.

The study specific log odds ratios per unit increase in exposure will be combined in a random effect model using the method of DerSimonian and Laird (31), with the estimate of heterogeneity being taken from the inverse-variance fixed-effect model.

Table 2. Approaches to derive missing information for meta-analyses in the CUP

Type of data	Problem	Approach
Dose-response data	Serving size is not quantified or ranges are missing, but group descriptions are given	Use serving size recommended in SLR
	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
	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 are missing	This information may be estimated by using the method of Chêne and Thompson(4;32) with a normal or lognormal distribution, as appropriate, or by taking midpoints (scaled in unbounded groups according to group numbers) if the number of groups is too small to calculate a distribution (3-4 groups)
Category data	Numbers of 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)

Publication and related bias (e.g. small study bias) will be explored through visual examination of funnel plots and Egger's test (33).

Heterogeneity between studies will be assessed visually from forest plots and with statistical tests (P value <0.05 will be considered statistically significant). Heterogeneity will be quantified with the I^2 statistic - where I^2 values of 25%, 50%, and 75% correspond to cut-off points for low, moderate, and high degrees of heterogeneity (34).

Stratified analyses will be performed to investigate potential sources of heterogeneity even if the initial overall test for heterogeneity is non-significant as these tests often have low power. The variables that will be explored as sources of heterogeneity are outcome definition, method of exposure assessment, gender, geographic area/country, level of adjustment (for instance, adjustment for dietary factors likely to be related to the risk of te investigated cancer), and in particular adjustment for *H.pylori* infection (for nongastric cardia and gastric cancer, site non-specified) , publication year, study size, length of follow-up. These variables will be explored if there are at least two studies in each of the categories of the variable. Meta-regression will be conducted when the number of studies allows it.

The interpretation 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 between the study characteristics and the size of associations.

9.7 Sensitivity analyses

Sensitivity analyses will be carried out to investigate how robust the overall findings of the CUP are relative to key decisions and assumptions that were made in the process of conducting the update. The purpose of doing sensitivity analyses is to strengthen the confidence that can be placed in the results.

Sensitivity analysis will be done as a minimum in the following cases:

- Including and excluding studies where there is some ambiguity as to whether they meet the inclusion criteria, for example it may be unclear what types of cancers are considered in a study (e.g. it is unclear if part of the cases might be of oesophageal cancer)
- Including and excluding studies where exposure was inferred by the authors (for example assigning a standard portion size when this is not provided) or other missing information was derived from the data.
- Influence-analyses where each individual study will be omitted in turn in order to investigate the sensitivity of the pooled estimates to inclusion or exclusion of particular studies(35)
- Including the results of pooling projects of cohort studies. In these analyses, the reviewer will check that studies in the pooled analyses are not included also as individual studies.

All analyses will be conducted in Stata/SE 12.1.

10. Reports

An updated report will be sent to the CUP Secretariat in 2014. The report will include the following elements:

10.1 Modifications of the approved protocol

Any modification required during the review will be described

10.2 Results of the search

Information on number of records downloaded, number of papers thought potentially relevant after reading titles and abstracts and number of papers included. The reasons for excluding papers should also be described.
This information will be summarised in a flowchart.

10. 3 Description of studies identified in the continuous update

Number of studies by study design and publication year.
Number of studies by population characteristics (gender, geographic area, others)
Number of studies by exposure (main heading and selected subheadings) and publication year
Number of studies by exposure and outcome subtype

10.4 Summary of number of studies by exposure and study type in the database, separated on studies identified in the continuous update and studies identified during the CUP.

10.5 Tabulation of study characteristics

The tables will include study characteristics (e.g. population, exposure, outcome, study design) and results of the study (e.g. direction and magnitude).

The tables will include the information required by the Panel to judge the quality of the studies included in the analyses (Newcastle –Ottawa quality assessment scale (36) for cohort studies and the Cochrane Collaboration's tool for assessing risk of bias (37).

Example of table of study characteristics (in two parts below):

Author, Year, country, WCRF Code	Study design	Country, Ethnicity, other characteristics	Age (mean)	Cases (n)	Non cases (n/person-years)	Case ascertainment	Follow-up (years)

Assessment details	Category of exposure	Subgroup	No cat	OR	(95% CI)	p trend	Adjustment factors						
							A	B	C	D	E	F	G

Where

A: Age
B: Ethnicity, race
C: Smoking
D: Anthropometric factors
E: Alcohol intake
F: Family history
G: Others, e.g. dietary factors, socioeconomic status, *H.pylori* infection

10. 6 Graphic presentation

Tabular presentation will be complemented with graphic displays when the number of studies justifies it. Study results will be displayed in forest plots showing relative risk estimates and 95% confidence interval of ''high versus low'' comparisons for each study. Dose-response graphs will be given for individual studies for which the information is available. Funnel plots will be shown when there are at least five studies.

10.7 Results of meta-analysis

Main characteristics of included and excluded studies in dose-response meta-analysis will be tabulated, and reasons for exclusions will be detailed.

The results of meta-analysis will be presented in tables and forest plots. The tables will include a comparison with the results of the meta-analyses undertaken during the SLR for the Second Expert Report.

All forest plots in the report will have the same format. Footnotes will provide quantified information (statistical tests and I^2 statistics) on the degree of heterogeneity between the displayed studies.

The results of meta-regression, stratified analyses and sensitivity analysis will be presented in tables and, when the number of studies justifies it, in forest plots.

Reference List

1. World Cancer Research Fund/ American Institute for Cancer Research: Food, Nutrition, Physical Activity, and the Prevention of Cancer: a global perspective. Washington DC: 2007.

2. World Cancer Research Fund/ American Institute for Cancer Research. Food, Nutrition and the Prevention of Cancer: a global perspective. Washington DC : 1997.

3. Romaguera D, Vergnaud AC, Peeters PH et al. Is concordance with World Cancer Research Fund/American Institute for Cancer Research guidelines for cancer prevention related to subsequent risk of cancer? Results from the EPIC study. Am J Clin Nutr 2012;96:150-63.

4. World Cancer Research Fund/ American Institute for Cancer Research. Systematic Literature Review. The SLR Specification Manual In . In: AICR, ed. Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective (Support Resource) . Washington DC: 2007.

5. Ferlay J, Shin HR, Bray F, Forman D, Mathers C and Parkin DM. GLOBOCAN 2008 v1.2, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. Lyon, France: International Agency for Research on Cancer; 2010. http://globocan iarc fr 2010.

6. Lauren P. The Two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt to a histo-clinical classification. Acta Pathol Microbiol Scand 1965;64:31-49.:31-49.

7. Gastric cancer and Helicobacter pylori: a combined analysis of 12 case control studies nested within prospective cohorts. Gut 2001;49:347-53.

8. de Vries AC, Kuipers EJ. Epidemiology of premalignant gastric lesions: implications for the development of screening and surveillance strategies. Helicobacter 2007;12 Suppl 2:22-31.:22-31.

9. Gonzalez CA, Agudo A. Carcinogenesis, prevention and early detection of gastric cancer: where we are and where we should go. Int J Cancer 2012;130:745-53.

10. de MC, Ferlay J, Franceschi S et al. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. Lancet Oncol 2012;13:607-15.

11. Monographs on the Evaluation of Carcinogenic Risks to Humans: Tobacco Smoke and Involuntary Smoking. (Vol 83). 2004. Lyon. IARC.

12. D'Elia L, Rossi G, Ippolito R, Cappuccio FP, Strazzullo P. Habitual salt intake and risk of gastric cancer: a meta-analysis of prospective studies. Clin Nutr 2012;31:489-98.

13. Dias-Neto M, Pintalhao M, Ferreira M, Lunet N. Salt intake and risk of gastric intestinal metaplasia: systematic review and meta-analysis. Nutr Cancer 2010;62:133-47.

14. Moskal A, Norat T, Ferrari P, Riboli E. Alcohol intake and colorectal cancer risk: a dose-response meta-analysis of published cohort studies. Int J Cancer 2007;120:664-71.

15. Druesne-Pecollo N, Latino-Martel P, Norat T et al. Beta-carotene supplementation and cancer risk: a systematic review and metaanalysis of randomized controlled trials. Int J Cancer 2010;127:172-84.

16. Saadatian-Elahi M, Norat T, Goudable J, Riboli E. Biomarkers of dietary fatty acid intake and the risk of breast cancer: a meta-analysis. Int J Cancer 2004;111:584-91.

17. Aune D, Greenwood DC, Chan DS et al. Body mass index, abdominal fatness and pancreatic cancer risk: a systematic review and non-linear dose-response meta-analysis of prospective studies. Ann Oncol 2012;23:843-52.

18. Aune D, Chan DS, Lau R et al. Carbohydrates, glycemic index, glycemic load, and colorectal cancer risk: a systematic review and meta-analysis of cohort studies. Cancer Causes Control 2012;23:521-35.

19. Aune D, Lau R, Chan DS et al. Dairy products and colorectal cancer risk: a systematic review and meta-analysis of cohort studies. Ann Oncol 2012;23:37-45.

20. Aune D, Chan DS, Vieira AR et al. Dietary compared with blood concentrations of carotenoids and breast cancer risk: a systematic review and meta-analysis of prospective studies. Am J Clin Nutr 2012;96:356-73.

21. Aune D, Chan DS, Greenwood DC et al. Dietary fiber and breast cancer risk: a systematic review and meta-analysis of prospective studies. Ann Oncol 2012;23:1394-402.

22. Aune D, Chan DS, Lau R et al. Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. BMJ 2011;343:d6617. doi: 10.1136/bmj.d6617.:d6617.

23. Aune D, Chan DS, Vieira AR et al. Dietary fructose, carbohydrates, glycemic indices and pancreatic cancer risk: a systematic review and meta-analysis of cohort studies. Ann Oncol 2012;23:2536-46.

24. Riboli E, Norat T. Epidemiologic evidence of the protective effect of fruit and vegetables on cancer risk. Am J Clin Nutr 2003;78:559S-69S.

25. Druesne-Pecollo N, Touvier M, Barrandon E et al. Excess body weight and second primary cancer risk after breast cancer: a systematic review and meta-analysis of prospective studies. Breast Cancer Res Treat 2012;135:647-54.

26. Kaiser Permanente SLR Team. Systematic Literature Review. The associations between food, nutrition and physical activity and the risk of endometrial cancer and underlying mechanisms. 2012. Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective (Support Resource).Washington DC: AICR , 2007 .

27. Jenab M, Slimani N, Bictash M, Ferrari P, Bingham SA. Biomarkers in nutritional epidemiology: applications, needs and new horizons. Hum Genet 2009;125:507-25.

28. Bekkering GE, Harris RJ, Thomas S et al. How much of the data published in observational studies of the association between diet and prostate or bladder cancer is usable for meta-analysis? Am J Epidemiol 2008;167:1017-26.

29. Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized dose-response data . Stata J 2006;6:40-57.

30. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. Am J Epidemiol 1992;135:1301-9.

31. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88.

32. Chene G, Thompson SG. Methods for summarizing the risk associations of quantitative variables in epidemiologic studies in a consistent form. Am J Epidemiol 1996;144:610-21.

33. Egger M, Davey SG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.

34. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.

35. Tobias A. Assessing the influence of a single study in meta-analysis. Stata Tech Bull 1999;47:15-7.

36. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 2010;25:603-5.

37. Cochrane Handbook for Systematic Reviews of Interventions . http://www.cochrane.org/training/cochrane-handbook . 2012.

Annex 1. WCRF - PUBMED SEARCH STRATEGY

1) Searching for all studies relating to food, nutrition and physical activity:

#1 diet therapy[MeSH Terms] OR nutrition[MeSH Terms]
#2 diet[tiab] OR diets[tiab] OR dietetic[tiab] OR dietary[tiab] OR eating[tiab] OR intake[tiab] OR nutrient*[tiab] OR nutrition[tiab] OR vegetarian*[tiab] OR vegan*[tiab] OR "seventh day adventist"[tiab] OR macrobiotic[tiab]
#3 food and beverages[MeSH Terms]
#4 food*[tiab] OR cereal*[tiab] OR grain*[tiab] OR granary[tiab] OR wholegrain[tiab] OR wholewheat[tiab] OR roots[tiab] OR plantain*[tiab] OR tuber[tiab] OR tubers[tiab] OR vegetable*[tiab] OR fruit*[tiab] OR pulses[tiab] OR beans[tiab] OR lentils[tiab] OR chickpeas[tiab] OR legume*[tiab] OR soy[tiab] OR soya[tiab] OR nut[tiab] OR nuts[tiab] OR peanut*[tiab] OR groundnut*[tiab] OR (seeds[tiab] and (diet*[tiab] OR food*[tiab])) OR meat[tiab] OR beef[tiab] OR pork[tiab] OR lamb[tiab] OR poultry[tiab] OR chicken[tiab] OR turkey[tiab] OR duck[tiab] OR fish[tiab] OR ((fat[tiab] OR fats[tiab] OR fatty[tiab]) AND (diet*[tiab] or food*[tiab] or adipose[tiab] or blood[tiab] or serum[tiab] or plasma[tiab])) OR egg[tiab] OR eggs[tiab] OR bread[tiab] OR (oils[tiab] AND and (diet*[tiab] or food*[tiab] or adipose[tiab] or blood[tiab]or serum[tiab] or plasma[tiab])) OR shellfish[tiab] OR seafood[tiab] OR sugar[tiab] OR syrup[tiab] OR dairy[tiab] OR milk[tiab] OR herbs[tiab] OR spices[tiab] OR chilli[tiab] OR chillis[tiab] OR pepper*[tiab] OR condiments[tiab] OR tomato*[tiab]
#5 fluid intake[tiab] OR water[tiab] OR drinks[tiab] OR drinking[tiab] OR tea[tiab] OR coffee[tiab] OR caffeine[tiab] OR juice[tiab] OR beer[tiab] OR spirits[tiab] OR liquor[tiab] OR wine[tiab] OR alcohol[tiab] OR alcoholic[tiab] OR beverage*[tiab] OR (ethanol[tiab] and (drink*[tiab] or intake[tiab] or consumption[tiab])) OR yerba mate[tiab] OR ilex paraguariensis[tiab]
#6 pesticides[MeSH Terms] OR fertilizers[MeSH Terms] OR "veterinary drugs"[MeSH Terms]
#7 pesticide*[tiab] OR herbicide*[tiab] OR DDT[tiab] OR fertiliser*[tiab] OR fertilizer*[tiab] OR organic[tiab] OR contaminants[tiab] OR contaminate*[tiab] OR veterinary drug*[tiab] OR polychlorinated dibenzofuran*[tiab] OR PCDF*[tiab] OR polychlorinated dibenzodioxin*[tiab] OR PCDD*[tiab] OR polychlorinated biphenyl*[tiab] OR PCB*[tiab] OR cadmium[tiab] OR arsenic[tiab] OR chlorinated hydrocarbon*[tiab] OR microbial contamination*[tiab]
#8 food preservation[MeSH Terms]
#9 mycotoxin*[tiab] OR aflatoxin*[tiab] OR pickled[tiab] OR bottled[tiab] OR bottling[tiab] OR canned[tiab] OR canning[tiab] OR vacuum pack*[tiab] OR refrigerate*[tiab] OR refrigeration[tiab] OR cured[tiab] OR smoked[tiab] OR preserved[tiab] OR preservatives[tiab] OR nitrosamine[tiab] OR hydrogenation[tiab] OR fortified[tiab] OR additive*[tiab] OR colouring*[tiab] OR coloring*[tiab] OR flavouring*[tiab] OR flavoring*[tiab] OR nitrates[tiab] OR nitrites[tiab] OR solvent[tiab] OR solvents[tiab] OR ferment*[tiab] OR processed[tiab] OR antioxidant*[tiab] OR genetic modif*[tiab] OR genetically modif*[tiab] OR vinyl chloride[tiab] OR packaging[tiab] OR labelling[tiab] OR phthalates[tiab]
#10 cookery[MeSH Terms]
#11 cooking[tiab] OR cooked[tiab] OR grill[tiab] OR grilled[tiab] OR fried[tiab] OR fry[tiab] OR roast[tiab] OR bake[tiab] OR baked[tiab] OR stewing[tiab] OR stewed[tiab] OR casserol*[tiab] OR broil[tiab] OR broiled[tiab] OR boiled[tiab] OR (microwave[tiab] and (diet*[tiab] or food*[tiab])) OR microwaved[tiab] OR re-heating[tiab] OR reheating[tiab] OR heating[tiab] OR re-heated[tiab] OR heated[tiab] OR poach[tiab] OR poached[tiab] OR steamed[tiab] OR barbecue*[tiab] OR chargrill*[tiab] OR heterocyclic amines[tiab] OR polycyclic aromatic hydrocarbons[tiab] OR dietary acrylamide[tiab]

#12 ((carbohydrates[MeSH Terms] OR proteins[MeSH Terms]) and (diet*[tiab] or food*[tiab])) OR sweetening agents[MeSH Terms]
#13 salt[tiab] OR salting[tiab] OR salted[tiab] OR fiber[tiab] OR fibre[tiab] OR polysaccharide*[tiab] OR starch[tiab] OR starchy[tiab] OR carbohydrate*[tiab] OR lipid*[tiab] OR ((linoleic acid*[tiab] OR sterols[tiab] OR stanols[tiab]) AND (diet*[tiab] or food*[tiab] or adipose [tiab] or blood[tiab] or serum[tiab] or plasma[tiab])) OR sugar*[tiab] OR sweetener*[tiab] OR saccharin*[tiab] OR aspartame[tiab] OR acesulfame[tiab] OR cyclamates[tiab] OR maltose[tiab] OR mannitol[tiab] OR sorbitol[tiab] OR sucrose[tiab] OR xylitol[tiab] OR cholesterol[tiab] OR protein[tiab] OR proteins[tiab] OR hydrogenated dietary oils[tiab] OR hydrogenated lard[tiab] OR hydrogenated oils[tiab]
#14 vitamins[MeSH Terms]
#15 supplements[tiab] OR supplement[tiab] OR vitamin*[tiab] OR retinol[tiab] OR carotenoid*[tiab] OR tocopherol[tiab] OR folate*[tiab] OR folic acid[tiab] OR methionine[tiab] OR riboflavin[tiab] OR thiamine[tiab] OR niacin[tiab] OR pyridoxine[tiab] OR cobalamin[tiab] OR mineral*[tiab] OR (sodium[tiab] AND (diet*[tiab] or food*[tiab])) OR iron[tiab] OR ((calcium[tiab] AND (diet*[tiab] or food*[tiab] or supplement*[tiab])) OR selenium[tiab] OR (iodine[tiab] AND and (diet*[tiab] or food*[tiab] or supplement*[tiab] or deficiency)) OR magnesium[tiab] OR potassium[tiab] OR zinc[tiab] OR copper[tiab] OR phosphorus[tiab] OR manganese[tiab] OR chromium[tiab] OR phytochemical[tiab] OR allium[tiab] OR isothiocyanate*[tiab] OR glucosinolate*[tiab] OR indoles[tiab] OR polyphenol*[tiab] OR phytoestrogen*[tiab] OR genistein[tiab] OR saponin*[tiab] OR coumarin*[tiab] OR lycopene[tiab]
#16 physical fitness[MeSH Terms] OR physical exertion[MeSH Terms] OR physical endurance[MeSH Terms] or walking[MeSH Terms]
#17 recreational activit*[tiab] OR household activit*[tiab] OR occupational activit*[tiab] OR physical activit*[tiab] OR physical inactivit*[tiab] OR exercise[tiab] OR exercising[tiab] OR energy intake[tiab] OR energy expenditure[tiab] OR energy balance[tiab] OR energy density[tiab]
#18 body weight [MeSH Terms] OR anthropometry[MeSH Terms] OR body composition[MeSH Terms] OR body constitution[MeSH Terms] OR obesity [MeSH Terms] OR obesity [MeSH Terms]
#19 weight loss[tiab] or weight gain[tiab] OR anthropometry[tiab] OR birth weight[tiab] OR birthweight[tiab] OR birth-weight[tiab] OR child development[tiab] OR height[tiab] OR body composition[tiab] OR body mass[tiab] OR BMI[tiab] OR obesity[tiab] OR obese[tiab] OR overweight[tiab] OR over-weight[tiab] OR over weight[tiab] OR skinfold measurement*[tiab] OR skinfold thickness[tiab] OR DEXA[tiab] OR bio-impedence[tiab] OR waist circumference[tiab] OR hip circumference[tiab] OR waist hip ratio*[tiab] OR weight change [tiab] OR adiposity [tiab] OR abdominal fat [tiab] OR body fat distribution [tiab] OR body size [tiab] OR waist-to-hip ratio [tiab]
#20 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19
#21 animal[MeSH Terms] NOT human[MeSH Terms]
#22 #20 NOT #21

2) Searching for all studies relating to stomach cancer:

#23 Stomach neoplasms[MeSH Terms]
#24 Stomach neoplasm*[tiab] OR stomach cancer*[tiab] OR stomach carcino* OR stomach tumo*[tiab] OR stomach metasta* [tiab] OR stomach malign*[tiab] OR stomach adenocarcinoma* [tiab]

#25 Gastric neoplasm* [tiab] OR gastric cancer*[tiab] OR gastric carcino* [tiab] or gastric tumo*[tiab] OR gastric metasta*[tiab] OR gastric malign*[tiab] OR gastric adenocarcinoma* [tiab]
#26 Gastrointestinal neoplasms[mesh terms] OR gastrointestinal neoplas*[tiab] OR gastrointestinal cancer*[tiab] OR gastrointestinal carcino*[tiab] OR gastrointestinal tumo*[tiab] OR gastrointestinal metasta*[tiab] OR gastrointestinal malign*[tiab] OR gastrointestinal adenocarcinoma*[tiab]
#27 Digestive tract neoplasm*[tiab] OR digestive tract cancer*[tiab] OR digestive tract carcino*[tiab] OR digestive tract tumo*[tiab] OR digestive tract metasta*[tiab] OR digestive tract malign*[tiab] OR digestive tract adenocarcinoma*[tiab]
#28 Alimentary tract neoplasm*[tiab] OR alimentary tract cancer*[tiab] OR alimentary tract carcino*[tiab] OR alimentary tract tumo*[tiab] OR alimentary tract metasta*[tiab] OR alimentary tract malign* OR alimentary tract adenocarcinoma*[tiab]
#29 Esophagogastric neoplasm*[tiab] OR esophagogastric cancer*[tiab] OR esophagogastric carcino* OR esophagogastric tumo*[tiab] OR esophagogastric metasta* [tiab] OR esophagogastric malign*[tiab] OR esophagogastric adenocarcinoma* [tiab] OR esophagogastric neoplasm*[tiab]
#30 Esophago gastric cancer*[tiab] OR esophago gastric carcino* OR esophago gastric tumo*[tiab] OR esophago gastric metasta* [tiab] OR esophago gastric malign*[tiab] OR esophago gastric adenocarcinoma* [tiab]
#31 Oesophagogastric neoplasm*[tiab] OR oesophagogastric cancer*[tiab] OR oesophagogastric carcino* OR oesophagogastric tumo*[tiab] OR oesophagogastric metasta* [tiab] OR oesophagogastric malign*[tiab] OR oesophagogastric adenocarcinoma* [tiab]
#32 Oesophago gastric neoplasm*[tiab] OR oesophago gastric cancer*[tiab] OR oesophago gastric carcino* OR oesophago gastric tumo*[tiab] OR oesophago gastric metasta* [tiab] OR oesophago gastric malign*[tiab] OR oesophago gastric adenocarcinoma* [tiab]
#33 Stomach adenoma*[tiab] OR gastric adenoma*[tiab] OR gastrointestinal adenoma*[tiab] OR digestive tract adenoma*[tiab] OR alimentary tract adenoma*[tiab] OR esophagogastric adenoma*[tiab] OR esophagogastric adenoma*[tiab] OR oesophagogastric adenoma*[tiab] OR oesophagogastric adenoma*[tiab]
#**34** #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33

3) Searching for all studies relating stomach cancer, and food, nutrition and physical activity:

#35 #22 AND #34

Annex 2. List of headings and exposure codes (minimum list)
**Indicated codes added during the CUP*

1 Patterns of diet

1.1 Regionally defined diets

*1.1.1 Mediterranean diet

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. diet diversity indexes) 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

Include here also age at first lactation, duration of breastfeeding, number of children breast-fed

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 diet diversity, meal frequency, frequency of snacking, dessert-eating and breakfast-eating should be reported here. Eating out of home should be reported here.

2 Foods

*2.0.1 Plant foods

2.1 Starchy foods

2.1.1 Cereals (grains)

* 2.1.1.0.1 Rice, pasta, noodles
* 2.1.1.0.2 Bread
* 2.1.1.0.3 Cereal

** Report under this subheading the cereals when it is not specified if they are wholegrain or refined cereals (e.g. fortified cereals)*

2.1.1.1 Wholegrain cereals and cereal products

* 2.1.1.1.1 Wholegrain rice, pasta, noodles
* 2.1.1.1.2 Wholegrain bread
* 2.1.1.1.3 Wholegrain cereal

2.1.1.2 Refined cereals and cereal products

* 2.1.1.2.1 Refined rice, pasta, noodles
* 2.1.1.2.2 Refined bread
* 2.1.1.2.3 Refined cereal

2.1.2 Starchy roots, tubers and plantains

* 2.1.2.1 Potatoes

2.1.3 Other starchy foods

**Report polenta under this heading*

2.2 Fruit and (non-starchy) vegetables

Results for "fruit and vegetables" and "fruits, vegetables and fruit juices" 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

*2.2.1.1.1 Carrots

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

*2.2.1.5.13 Tomatoes
*2.2.1.5.1 Fresh beans (e.g. string beans, French beans) and peas

Other non-starchy vegetables' should include foods that are botanically fruits but are eaten as vegetables, e.g. 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. 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

*2.2.2.0.1 Fruit, dried
*2.2.2.0.2 Fruit, canned
*2.2.2.0.3 Fruit, cooked

2.2.2.1 Citrus fruit

2.2.2.1.1 Oranges
2.2.2.1.2 Other citrus fruits (e.g. grapefruits)

2.2.2.2 Other fruits

*2.2.2.2.1 Bananas
*2.2.2.2.4 Melon
*2.2.2.2.5 Papaya
*2.2.2.2.7 Blueberries, strawberries and other berries
*2.2.2.2.8 Apples, pears
*2.2.2.2.10 Peaches, apricots, plums
*2.2.2.2.11 Grapes

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)

*2.3.1 Soya, soya products

*2.3.1.1 Miso, soya paste soup
*2.3.1.2 Soya juice
*2.3.1.4 Soya milk
*2.3.1.5 Tofu

*2.3.2 Dried beans, chickpeas, lentiles
*2.3.4 Peanuts, peanut products

Where results are available for a specific pulse/legume, 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

*2.5.1.2.1 Ham
*2.5.1.2.1.7 Burgers
*2.5.1.2.8 Bacon
*2.5.1.2.9 Hot dogs
*2.5.1.2.10 Sausages

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

*2.5.1.3.1 Beef
*2.5.1.3.2 Lamb
*2.5.1.3.3 Pork
*2.5.1.3.6 Horse, rabbit, wild meat (game)

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.

*2.5.1.5 Offals, offal products (organ meats)

2.5.2 Fish

*2.5.2.3 Fish, processed (dried, salted, smoked)
*2.5.2.5 Fatty Fish
*2.5.2.7 Dried Fish
*2.5.2.9 White fish, lean fish

2.5.3 Shellfish and other seafood

2.5.4 Eggs

2.6 Fats, oils and sugars

2.6.1 Animal fats

*2.6.1.1 Butter
*2.6.1.2 Lard
*2.6.1.3 Gravy
*2.6.1.4 Fish oil

2.6.2 Plant oils
2.6.3 Hydrogenated fats and oils

*2.6.3.1 Margarine

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

*2.7.1 Milk, fresh milk, dried milk

*2.7.1.1 Whole milk, full-fat milks
*2.7.1.2 Semi skimmed milk, skimmed milk, low fat milk, 2% Milk

*2.7.2 Cheese

*2.7.2.1 Cottage cheese
*2.7.2.2 Cheese, low fat

*2.7.3 Yoghurt, buttermilk, sour milk, fermented milk drinks

*2.7.3.1 Fermented whole milk
*2.7.3.2 Fermented skimmed milk

*2.7.7 Ice cream

2.8 Herbs, spices, condiments

*2.8.1 Ginseng
*2.8.2 Chili pepper, green chili pepper, red chili pepper

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.

*2.9.1 Cakes, biscuits and pastry
*2.9.2 Cookies
*2.9.3 Confectionery
*2.9.4 Soups
*2.9.5 Pizza
*2.9.6 Chocolate, candy bars
*2.9.7 Snacks

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 and vegetable juices

*3.5.1 Citrus fruit juice
*3.5.2 Fruit juice
*3.5.3 Vegetable juice
*3.5.4 Tomato juice

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-*nitrososamines 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.

1 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

*5.1.5 Glycemic index, glycemic load

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. Past alcohol refers, for example, to intake at age 18, during adolescence, etc.

*5.4.1 Total Alcohol (as ethanol)

*5.4.1.1 Alcohol (as ethanol) from beer
*5.4.1.2 Alcohol (as ethanol) from wine
*5.4.1.3 Alcohol (as ethanol) from spirits
*5.4.1.4 Alcohol (as ethanol) from other alcoholic drinks
* 5.4.1.5 Total alcohol (as ethanol), lifetime exposure

* 5.4.1.6 Total alcohol (as ethanol), past

5.5 Vitamins

*5.5.0 Vitamin supplements
*5.5.0.1 Vitamin and mineral supplements
*5.5.0.2 Vitamin B supplement

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

*5.5.3.1 Total folate
*5.5.3.2 Dietary folate
*5.5.3.3 Folate from supplements

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

*6.1.2.1 Frequency of occupational physical activity
*6.1.2.2 Frequency of recreational physical activity

6.1.3 Intensity of physical activity

*6.1.3.1 Intensity of occupational physical activity
*6.1.3.2 Intensity of recreational physical activity

6.1.4 Duration of physical activity

*6.1.4.1Duration of occupational physical activity
*6.1.4.2Duration of recreational physical activity

6.2 Physical inactivity

6.3 Surrogate markers for physical activity e.g. occupation

7 Energy balance

7.1 Energy intake

*7.1.0.1 Energy from fats
*7.1.0.2 Energy from protein
*7.1.0.3 Energy from carbohydrates
*7.1.0.4 Energy from alcohol
*7.1.0.5 Energy from all other sources

7.1.1 Energy density of diet

7.2 Energy expenditure

1.1.1 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

Annex 3. Tables of excluded and included 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 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.

Measured in	Include	Exclude
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* [glycitein, O-desmethylangolensin, equol, enterodiol, and enterolactone] 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
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
Measured	**Include**	**Exclude**

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

Exposure	Measured in	Valid?	Reason (Willett)	Reason (Margetts / Nelson)
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	Yes (p 197)

			sources (p 193)	
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)
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

Exposure	Measured in	Valid?	Reason (Willett)	Reason (Margetts / Nelson)
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)
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	Yes (p 193). Relationship between selenium intake and biomarkers is reasonably good. Urine: reasonable marker, plasma reflects intake provided that

			(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)	the range of variation is large. Red cell and glutathione perioxidase are markers of longer-term intakes. Hair and toenails are alternative possibilities, although contamination of hair samples with shampoo must be controlled for
Exposure	**Measured in**	**Valid?**	**Reason (Willett)**	**Reason (Margetts / Nelson)**
Glutathione perioxidase	Plasma Serum Erythrocytes Blood	No	Is poor measure of selenium intake among persons with moderate and high exposure (p 206)	
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	
Monounsat 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	
Exposure	**Measured**	**Valid?**	**Reason (Willett)**	**Reason (Margetts /**

	in			**Nelson)**
Polyunsat fatty acids	Adipose tissue	Yes	Yes (p 220)	No info
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=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 Urinary	Validity Validity	0.37 -	Correlation between B6 and plasma pyridoxal phosphate levels in 280 healthy men =0.37 (Willett p203)
Folacin	Serum Erythrocyte	Validity	0.56 0.51	Correlation of 0.56 in Framington Heart study 385 subjects (serum) Correlation in 19 elderly subjects (erythrocyte) (Willet p204)
Magnesium	Serum	Validity	0.27	Correlation between intake with supplements 0.27 in 139 men and 0.15 without supplements (Willett p211)

Nutrient	**Biologic tissue**	**Val./reproduc**	**Coef**	**Details**
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)
Nutrient	**Biologic tissue**	**Val./reproduc**	**Coef**	**Details**
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)
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)
Nutrient	**Biologic tissue**	**Val./reproduc**	**Coef**	**Details**
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

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