Kidney Cancer Protocol *Version 1*

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

Prepared by: CUP team, Imperial College London

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.

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 ongoing review of nutrition and cancer research on food, nutrition, physical activity and body fatness, and cancer risk.

The WCRF/AICR second expert report was informed by a process of systematic literature reviews (SLRs) all of the evidence published. The CUP follows a similar scientific process to ensure a consistent approach to reviewing the evidence. To improve cost-efficiency of the process, the CUP will systematically review the evidence from cohort studies and randomised controlled trials, the two study designs on top of the hierarchy of evidence [3]. The review will be complemented by a narrative review of pooled analyses of large consortia of case-control studies.

WCRF/AICR has convened an independent panel of experts (the CUP Panel) consisting of leading scientists in the field of diet, physical activity, obesity and cancer who will consider the results of the systematic literature reviews and meta-analyses produced by the review team at Imperial College, and will draw conclusions before making recommendations. The CUP 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.

The continuous update of the epidemiological evidence on food, nutrition, physical activity and the risk of kidney cancer will be conducted by a team of scientists at ICL. The peer-reviewed protocol will represent the agreed plan for the continuous update project. Should departure from the agreed plan be considered necessary at a later stage, this must be agreed by the CUP Expert Panel and the reasons documented.

An independent panel of experts is responsible for reviewing the CUP findings and for making judgements and recommendations based on the body of the scientific evidence.

Background

Renal (kidney) cancer

Cancer of the kidney is the 11th most common cancer site in men (a*ge-standardised rate per 100,000 = 5.2)* and the 15th most common cancer site for women (a*ge-standardised rate per 100,000=2.8)*[4] (**Figure 1**) [4].
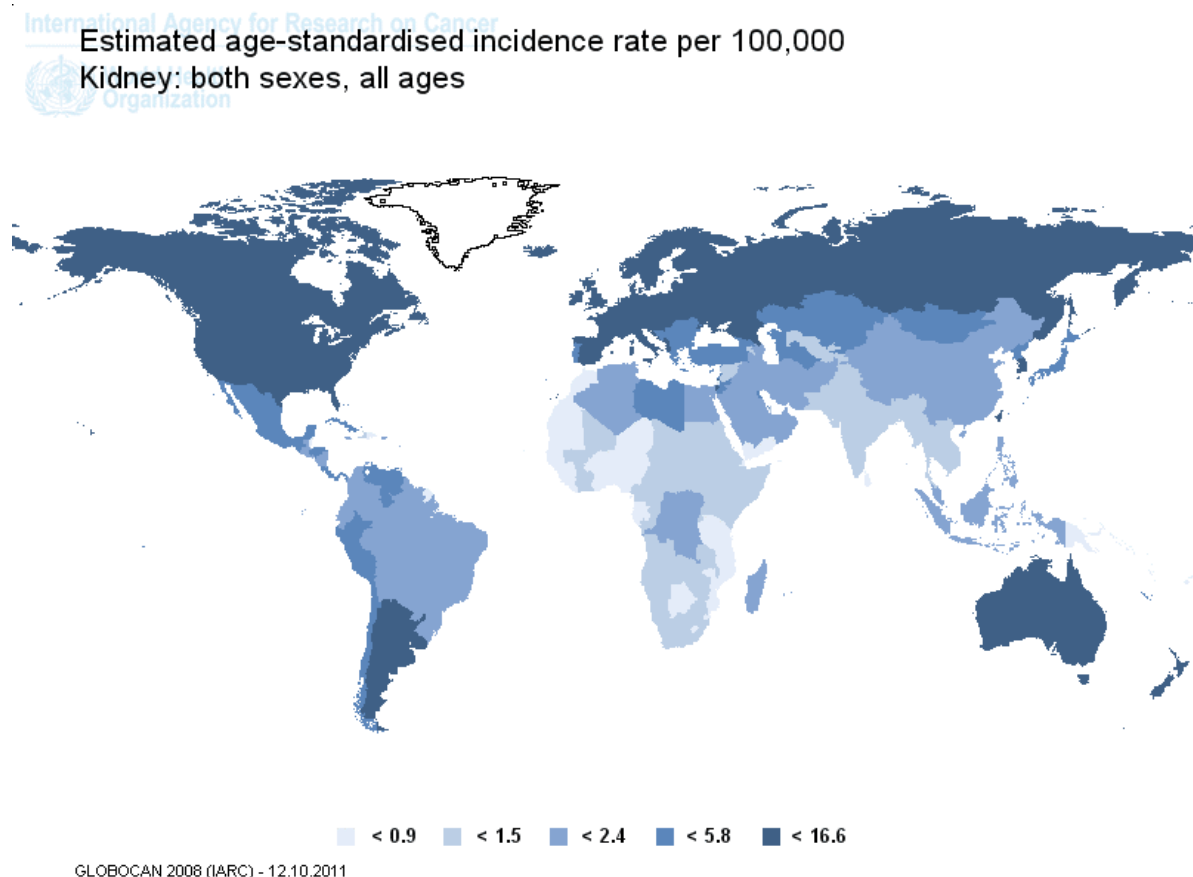
Figure 1. Estimated age (world)-standardised incidence and mortality rate of cancer per 100 000. World. 2008 [4]

GLOBOCAN 2008 (IARC) (12.10.2011)
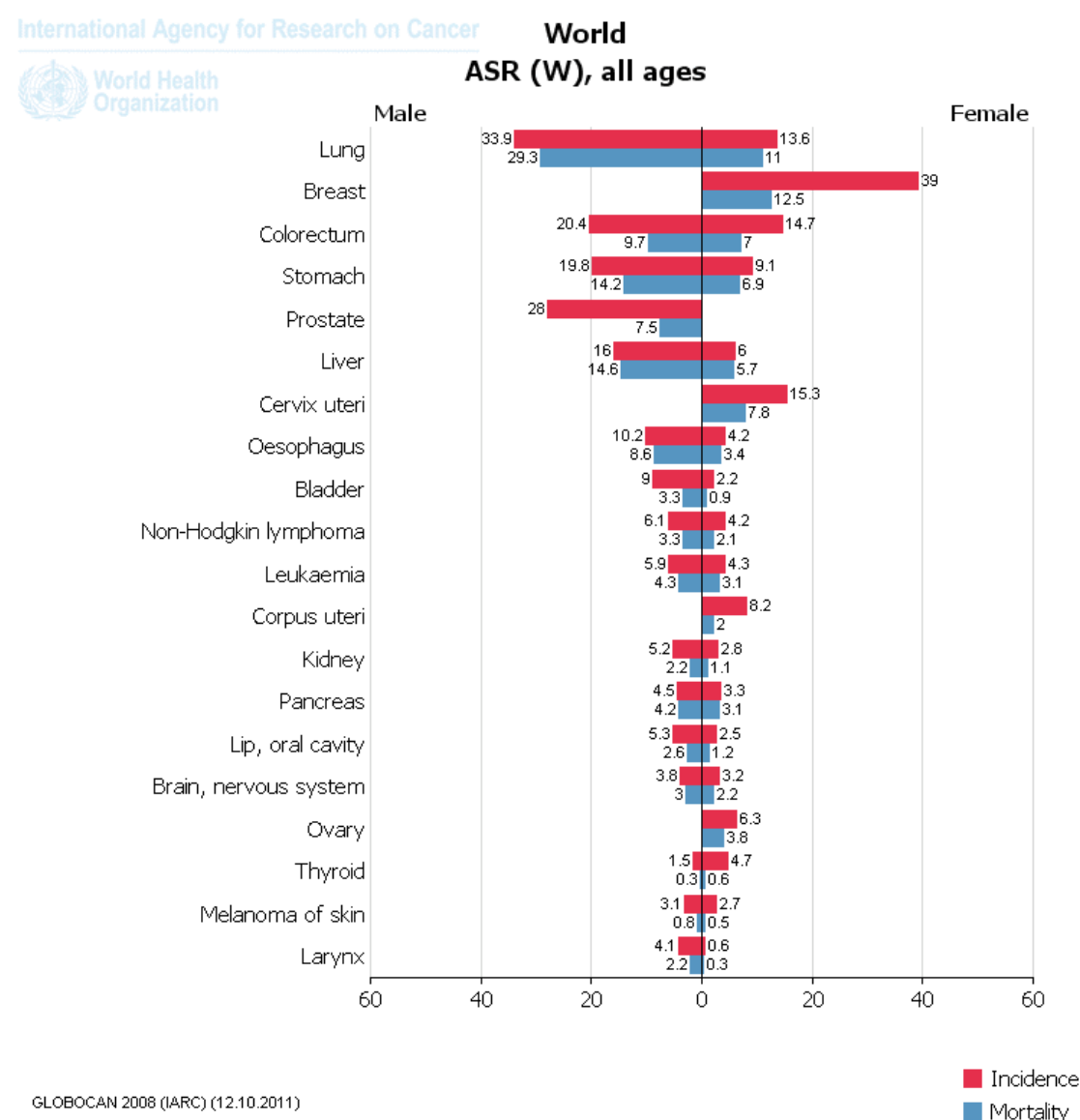
The incidence of kidney cancer rises with age, peaking between age 55 years and 75 years, and is about two times more common in men than in women[4]. There is a considerable geographic variation in incidence, with relatively high rates in North America, Europe, some Latin American countries and Australia (**Figure 2**). Incidence rates are also high in parts of Asia, and the Middle East, whereas rates are low in

2

India and Sub-Saharan Africa. The countries with the highest incidence of kidney cancer for both sexes are Belarus, Croatia and Israel[4]. The highest estimated mortality is in Czech Republic, which might be related to smoking habits, among other, not yet clarified, risk factors [4].

Figure 2. Estimated age-standardised incidence of kidney cancer per 100 000. World 2008 [4]

Estimated age-standardised incidence rate per 100,000
Kidney: both sexes, all ages

< 0.9   < 1.5   < 2.4   < 5.8   < 16.6

GLOBOCAN 2008 (IARC) - 12.10.2011

The most frequent type of kidney cancer is renal parenchyma cancer, mainly of the adenocarcinoma cell type (renal cell carcinoma), representing 80- 90% of all primary renal malignant tumours. The remainder 10% accounts for renal pelvis cancer, mostly from the transitional cell type [5,6].

Known risk factors for kidney cancer are cigarette smoking (male smokers are at an approximately 50% increased risk of developing this cancer compared with never smokers, while women smokers have 20% increase risk[5]), obesity, and personal history of hypertension and diabetes [5-8].

In the past, evidence on dietary factors was rather limited. In the judgment of the Panel of the WCRF-AICR second expert report[2] on kidney cancer, there was available convincing evidence that high body fatness increase the risk of kidney cancer. The evidence on dietary and nutritional, and physical activity factors examined was inconclusive (**Figure 3**).

More recently, the results of pooled studies and meta-analyses concur with the WCRF-AICR second expert report[2] conclusions that obesity increases the risk of

kidney cancer in men and women[9,10]. In addition, a pooled study on fruit and vegetable consumption reported that an increasing intake of these foods is associated with declining risk of renal cell cancer [11]. Other pooled analyses showed inconclusive data on meat and processed meat intake and the association with kidney cancer risk[12,13]. A pooled analysis on coffee consumption showed a modestly lower risk of renal cell cancer[14]. The same analysis concluded that tea also have an inverse association with renal cell cancer risk[14].

Figure 3. Summary of judgements of the 2007 Second Expert Report on kidney cancer [2]

FOOD, NUTRITION, PHYSICAL ACTIVITY, AND CANCER OF THE KIDNEY

In the judgement of the Panel, the factors listed below modify the risk of cancer of the kidney. Judgements are graded according to the strength of the evidence.

	DECREASES RISK	INCREASES RISK
Convincing		**Body fatness**
Probable		
Limited — suggestive		Arsenic in drinking water[1]
Limited — no conclusion	Cereals (grains) and their products; vegetables; fruits; meat; poultry; fish; eggs; milk and dairy products; total fat; soft drinks; tea; alcoholic drinks (protective effect)[2]; carbohydrate; protein; vitamin A; retinol; vitamin C; vitamin E; beta-carotene; flavonol; Seventh-day Adventist diets; physical activity; body fatness at age 18–20; weight at age 18–20; birth weight; adult attained height; age at menarche; energy intake.	
Substantial effect on risk unlikely	Coffee; alcoholic drinks (adverse effect)[2]	

1   The International Agency for Research on Cancer has graded arsenic and arsenic compounds as Class 1 carcinogens. The grading for this entry applies specifically to inorganic arsenic in drinking water.
2   The evidence was sufficient to judge that alcoholic drinks were unlikely to have an *adverse* effect on the risk of kidney cancer; but it was inadequate to draw a conclusion regarding a protective effect.

For an explanation of all the terms used in the matrix, please see chapter 3.5.1, the text of this section, and the glossary.

World Cancer Research Fund   American Institute for Cancer Research

1. Research question

The research topic of the continuous update of the epidemiological evidence on food, nutrition, physical activity and the risk of kidney cancer is:

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

The main objective is:

To summarize the evidence from prospective studies and clinical trials on the association between food, nutrition, physical activity, overweight and obesity with the risk of kidney cancer in men and women.

2. Review team

Name	Current position at IC	Role within team
Teresa Norat	Principal Research Fellow	Principal investigator
Rui Vieira	Data manager	Responsible of the data management, the design and architecture of the database
Doris Chan	Research Assistant	Nutritional epidemiologist, supervisor of data entry, analyst
Dagfinn Aune	Research Assistant	Nutritional epidemiologist, supervisor of data entry, analyst
Ana Rita Vieira	Research Assistant	Nutritional epidemiologist, reviewer
Deborah Navarro	Research Assistant	Nutritional epidemiologist, reviewer

Review coordinator, WCRF: Rachel Thompson

Statistical advisor: Darren Greenwood, senior Research Lecturer, University of Leeds

3. Timeline

The SLR's for the Second Expert Report ended in December 30th 2005. The continuous update will include the articles from prospective studies and clinical trials published from January 1st 2006. The reviewers will verify that there are not duplicities in the database. With that purpose, a module for article search has been implemented in the interface for data entry.

Meta-analyses and pooled analyses of case-control and cohort studies will be identified in the literature search and used as support for interpretation of results.

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

Task	Deadline
Start Medline search of relevant articles published from January 2006	1st December, 2011
Review abstracts and citations identified in initial electronic search. Select papers for complete review	Monthly
Review relevant papers. Select papers for data extraction	Monthly
Data extraction	Monthly
Start quantitative analysis	January 2013
Start report writing	April 2013
Send report to WCRF-AICR	June 2013
Transfer Endnote files to WCRF	June 2013

*For the report to WCRF-AICR, search will end in December 31st 2012

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 used PubMed as platform. Data provided by the SLR's [15] for the Second Expert Report suggest that searching in databases other than Medline was not cost effective.

4.2. Hand searching for cited references

To check for potential missed articles, the review team will hand search the references of reviews and meta-analyses identified during the search, as well as the references of the articles relevant to the review published in the last two years before the preparation of the CUP report.

If the hand searching shows that articles have been missed, the CUP review team will consider other strategies, such as implementing searches in other databases (e.g. Embase, Ovid, Cochrane Library).

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[16]. The full search strategy for food, nutrition and physical activity is in Annex 1.

a) **#1 to #22** Searching for studies relating to food, nutrition and physical activity (Appendix 1)

b) **Searching for all studies relating to kidney cancer:**
  c) #23 renal neoplasma [MeSH terms] OR kidney neoplasma [MeSH terms] OR urethral neoplasma [MeSH terms]

d) # 24 kidney OR renal OR urethral AND (malign* [tiab] OR cancer*[tiab] OR carcinoma*[tiab] OR tumor*[tiab] OR tumour*[tiab] OR adenocarcinoma* OR sarcoma)
e) # 25 hypernephroma* OR grawitz tumo* OR wilms tumo* OR nephroma*
f)
g) #26: # 23 OR #24 OR #25

5. Study selection criteria

5.1 Inclusion criteria

The articles to be included in the review:

- Have to present results on at least one of the exposures as listed in the section 20 of the SLR specification manual (Annex 2).
- Must have as outcome of interest incidence or mortality from: kidney cancer (not specified), renal parenchyma cancer and/or renal pelvis cancer.

- Have to present results from an epidemiologic study in men or women of one of the following types[†]:

  o Randomized controlled trial
  o Group randomized controlled trial (Community trial)
  o Prospective cohort study
  o Nested case-control study
  o Case-cohort study
  o Historical cohort study

- Have to be included in *PubMed* from January 1st 2006 (closure date of the database for the Second Expert Report[16]).

† Pooled analysis of cohort studies and case-control consortia will be identified in the search, but their results will not be included in the meta-analyses. They will be used for narrative reviews and as support in the interpretation of the data. The rationale is that the inclusion of a pooled result as a single study in a meta-analysis may apparently decrease the heterogeneity, if included as a single study. Therefore, if study-specific results are shown in the manuscript of a pooled analysis of cohort studies, the cohort-specific results will be extracted and included in the meta-analyses in the CUP. Additionally sensitivity analyses will be further conducted to test the effect of the inclusion of the results of pooling projects in the overall estimate. Special care will be taken to avoid study duplicities.
Filters for study design will not be implemented in the search strategy.

5.2 Exclusion criteria

The articles to be excluded from the review:

- Do not report measure of association between any of the relevant exposures and outcomes.
- Focus on Wilm's tumour or cancer related to schistosomiasis infection.
- Cohort studies in which the only measure of the relationship between the relevant exposure and outcome is the mean difference of exposure, because the difference is not adjusted for main confounders.
- Are supplement to the main manuscript (e.g. Authors' Reply).
- Published abstracts

6. Article selection

All references obtained with the search in PubMed will be archived in Reference Manager Databases. The variables in the Reference manager files will be those generated using the filter Medline for importing data. Additionally, customized fields will be implemented.

The article selection will follow three steps:

1. An electronic search will be undertaken within Reference Manager to identify irrelevant records. This will be achieved by applying a list of stop words developed and tested during the preparation of the WCRF-AICR Second Expert Report. The list of stop words was compiled from terms that describe surgical, diagnostic or oncology procedures, animal and in vitro studies.

Stop Words for use within Reference Manager Database

Resection
Radiotherapy
Radiochemotherapy
Cisplatin
Fluorouracil
5 FU
Gemcitabine
Mitomycin C
Carboplatin
Methotrexate
Vinblastine
Gemcitabine
Doxorubicin
Antineoplastic
Peptides
Cell
Inhibitor
Novel
Model
Receptor
Antibody
P53
Transgenic
Mice
Hamster
Rat
Dog

Cat
In vitro

2. In a second step the titles and abstracts of the remaining articles will be assessed by two reviewers using the inclusion criteria. 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 of all studies that are not clearly ineligible will then be obtained and the two reviewers will assess all obtained papers.

Disagreements between the reviewers will be resolved by discussion with the principal investigator. If a retrieved paper reports outcomes for more than one cancer site, the principal investigator will be informed and will in turn inform the teams performing other reviews in the CUP.

6.1 Reference Manager Files

Three Reference Manager files will be created:

1) A file containing the results of the initial search. The study identifier should be entered under a customized field titled 'label'. Another customised field named 'inclusion' should be marked 'in' or 'out' for each paper, thereby indicating which papers were deemed potentially relevant based on an assessment of the title and abstract.
2) A file containing the excluded papers. The study identifier should be entered under a customized field titled 'label'. Another customised field named 'reasons' should include the reason for exclusion for each paper. This file will be named Kidney-excluded.
3) A file containing the relevant papers. The study identifier should be entered under a customized field titled 'label'. Another customised field named "study design" should include a letter representing the study design of each paper, allocated using the study design algorithm (in **Figure 5**). This file will be named Kidney- included.

The Reference Management databases will be converted to EndNote and sent once per year to the WCRF Secretariat.

7. Data extraction

The IC team will update the WCRF-AICR central database using the interface created at Imperial College for this purpose. One of the screens of the interface for data extraction is shown in **Figure 4.** Data extraction will include study characteristics that are potential sources of heterogeneity, such as study design, study population, country, methods of exposure assessment, cancer subtype, analytical methods, adjustment variables, matching variables, whether methods for correction of measurement error were used, study size, length of follow up, and study results will be extracted. Results related to gene nutrient interactions and interactions with other potential modifiers will also be extracted.

.

Figure 4. Screen of 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 5**.

7.2 Study identifier

The CUP team will use the same labelling of articles used in the SLR process for the Second Expert Report[2]: the unique identifier for an article will be constructed using a 3-letter code to represent the cancer site: KID (kidney cancer), followed by a 5-digit number that will be allocated in sequence.

Figure 5. Study design algorithm (From: SLR specification manual)

SLR specification manual - version 15
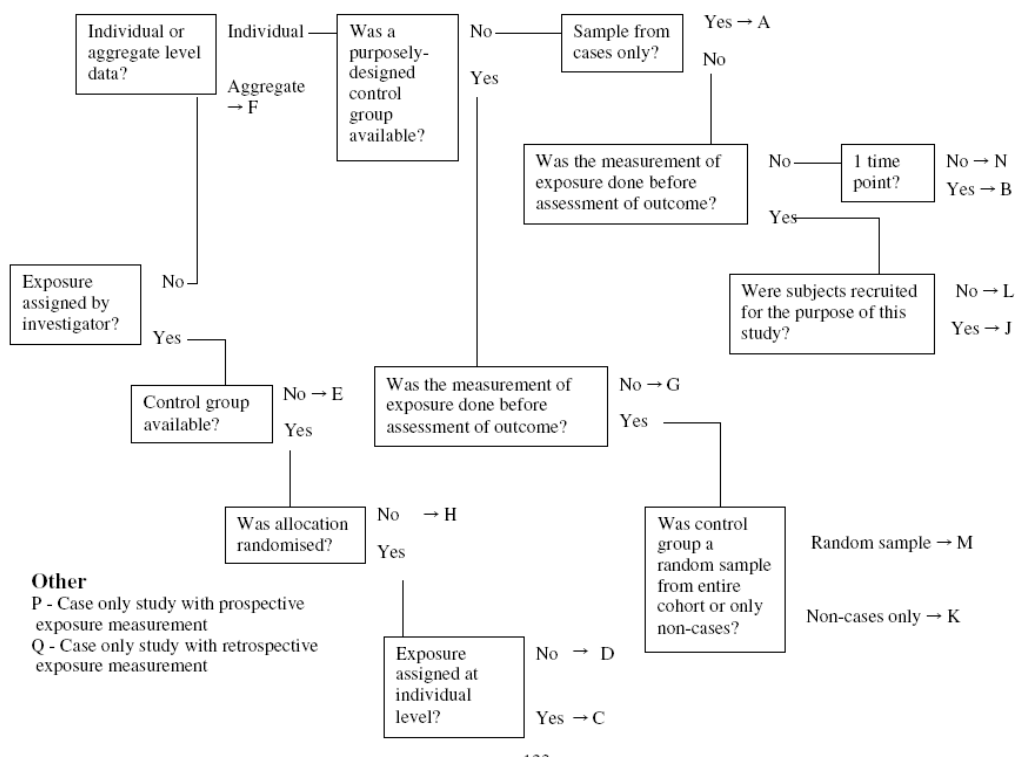
Key to study design algorithm
Study design A Case-study / case series
Study design B Cross-sectional study
Study design C Randomised controlled trial
Study design D Group randomised control trial
Study design E Uncontrolled trial
Study design F Ecologic study
Study design G Case-control study
Study design H Non-randomised 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

7.3 Codification of exposures

The CUP will use the exposure and sub-exposure labels and codes listed in the SLR Guidelines[2] for the Second Expert Report. Additional codes for sub-exposures were added during the SLRs for the Second Expert Report and during the CUP at Imperial College. The exposures listed represent the minimum list of exposures to be examined. These exposures are programmed in the interface for data entry generated at Imperial College with the purpose of facilitating data entry.

The updated list of selected codes for exposures is in Annex 2.

Main exposure levels and codes 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 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 foetal 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, the CUP team will follow the conservative approach of extracting the data for all biomarkers reported in

the relevant studies, independently of whether validity and reliability had been or not fully documented.

The "biomarkers" whose validity was considered not yet been fully documented are:

Vit D: 1.25 $(OH)_2D$, Alkaline phosphatase activity (serum)

Iron (serum, hair, nails)

Copper (plasma, serum, hair)

Glutathione peroxidase (plasma, serum, erythrocytes, blood)

Zinc, metallotein levels (any)

Lipids: total fats (any)

Cholesterol, LDL (any)

Lipoprotein levels (serum)

Monounsaturated fatty acids (oleic acid) (plasma, adipose tissue)

Saturated fatty acids (palmitic acid, stearic acids) (plasma)

Protein (any)

The full list of biomarkers for which the evidence on validity and reliability was considered strong for the purpose of the Second Expert Report and the rationale is in Annex 3.

7. 4 Labelling of Results

The results for each exposure obtained using all the models reported in the paper and all the subgroup or stratified analyses will be extracted by the reviewer.

The reviewer should label the results as unadjusted, less adjusted, intermediately adjusted, most adjusted model.

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, except when the most adjusted model is a "mechanistic" model, that is a model that includes variables likely to be in the causal pathway (e.g. milk intake as main exposure in a model adjusted for dietary calcium). When such results are reported, the "intermediately" adjusted result with the highest number of covariates will be indicated as "best model".

Sometimes, potential risk factors are not kept in the final model because their inclusion does not substantially modify the risk estimates. If this is specified in the article text, this model should also 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 quantile ranges are provided over where missing) and in certain circumstances, the number of cases (highest).

8. Quality control

All the data extracted during the first year of the continuous update will be checked by a second reviewer at Imperial College. In the second year, an initial 10% random sample of the data extracted will be assessed by a second reviewer. If errors are identified, another 10% will be assessed by a second reviewer. The process will be repeated until no error is identified. The purpose of implementing this procedure is to be more cost-effective. This procedure will be followed only for the reviewers working in the CUP of kidney cancer for more than one year.

9. Data analysis

The overall aim of data synthesis is to collate and summarise the results of the studies included in the CUP. Meta-analytic and narrative aspects of the data analysis will complement each other.

The primary analyses will include studies that explicitly report on renal parenchyma cancer or renal cell (adeno)carcinoma or studies that do not distinguish type of kidney cancer. Studies on cancer of the renal pelvis and/or cancer of the ureter and/or transitional cell carcinoma will be analysed separately.

Summary estimates will be prepared for trials (if identified) and studies based in cohorts (including case-cohort analyses and nested case-control studies) separately. Studies with mortality as outcome will be analysed separately.

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 from each study will be pooled first using fixed effect models and then included in the meta-analysis of "Both gender". This is essentially equivalent to including the overall estimate and will provide a better estimate of heterogeneity across studies.

Where results from two or more cohort studies are included in the same paper, the study results will be included separately and the pooled result will not be included. This maintains the independence of observations included and permits to look at heterogeneity across study results. The inclusion of results of the pooled analysis will be examined in sensitive analyses (see Section 9.5).

In a forest plot, the studies will be ordered by publication year.

9.1 Statistical Methods

For each study where this is possible we will derive estimates of the log odds ratio per unit increase in exposure and their standard errors using the method of Greenland and

Longnecker [17]. This method accounts for the correlation between relative risks estimates with respect to the same reference category. We will use the "best" (most adjusted risk estimate) from each study and if no model is considered the "best", we will use the most adjusted model that is not mechanistic model. 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, with the estimate of heterogeneity being taken from the inverse-variance fixed-effect model [18].

The analyses will be conducted using STATA version 11.1 (College Station, TX, USA).

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

A dose-response meta-analysis for a particular exposure and outcome will be conducted in the CUP when at least two trials or two cohort studies are identified that were not included in the meta-analyses conducted during the SLR for the Second Expert Report. These refer to studies providing enough information to conduct dose-response meta-analysis.

The dose-response meta-analysis will include study results identified during the SLR and study results identified during the CUP. Special care will be taken to avoid including more than once the results of the same study. Where more than one result on the same exposure and outcome has been published from a particular study, the result using the larger number of cases for analysis will be selected. This is often retrieved in the most recent paper.

9.3   Selection of results 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.
In addition to effect measures and their standard errors (or confidence intervals), other elements required to derive the dose-response slopes in each study are: (i) number of individuals with the disease for each exposure category (ii) person-years -or number of individuals without the disease in case-control analysis- for each exposure category, and (iii) exact cut-offs of exposure categories, or mean or median of each category. The exposure should be expressed in the same units. When the effect measure per unit of increase is reported in the articles, this effect measure will be used directly in the dose-response meta-analysis.

However, reporting in published research is often incomplete and this may result in exclusion of results from meta-analyses. 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. It has been reported that only 64% of the results of cohort studies on kidney and prostate cancer provided enough data to be included in dose-response meta-analysis in the SLR for the Second Expert Report[16]. Moreover, results that showed evidence of an association were more likely to be usable in dose-response meta-analysis than results that found no such evidence.

9.4 Derivation of data required for meta-analyses.

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[19]. The approaches are summarized in **Table 1**.

Table 1. 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[13]
	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 [20] 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)

9.5 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. If the sensitivity analyses results do change in a way that might lead to different conclusions, this indicates a need for greater caution in interpreting the results.

Sensitivity analysis will be done as a minimum for:
• Including and excluding studies where there is some ambiguity as to whether they meet the inclusion criteria, for example for outcome type (e.g. kidney cancer non specified).
• Including and excluding studies where exposure unit or other missing information was inferred by the reviewers.

- Including and excluding study results that were not adjusted for body mass index (except analysis on body fatness) or smoking
- Influence-analyses where each individual study will be omitted in turn will be done to investigate the sensitivity of the pooled estimates to inclusion or exclusion of particular studies[21]
- Including the results of pooling projects of cohort studies or large case-control studies.

During the analyses, when the number of studies makes it possible, the IC team will conduct sensitivity analyses using as criteria, those included in the Newcastle – Ottawa quality assessment scale[22]. However, scoring of cohort study quality will not be used as it is unclear which of the many published scales is better. For clinical trials –if any is identified in the search- the CUP team will use The Cochrane Collaboration's tool for assessing risk of bias[23] and will explore in stratified analysis whether study quality score influences the results.

9.6 Analysis of heterogeneity and potential bias

Small study bias (e.g. publication bias) was explored through visual examination of funnel plots and Egger's test.

Heterogeneity between studies will be assessed visually from forest plots, with statistical tests and 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[24].

Meta-regression and 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 examined as sources of heterogeneity where possible include outcome definition, gender, geographic area/country, level of adjustment, publication year, study size, length of follow-up.

Other variables that may be considered as source of heterogeneity are characterisation of the exposure (FFQ, recall, diary, anthropometry etc.), and exposure range (including correction for measurement error, length of intervention).

The interpretation of the exploration of heterogeneity 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.

Where heterogeneity not explicable by chance is detected and substantial variability exists between studies, it will be considered whether it is appropriate to present a combined estimate for all studies.

10. Reports

An updated report will be produced in 2013. 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

Information on the characteristics (e.g. population, exposure, outcome, study design)
and results of the study (e.g. direction and magnitude) will be summarised in tables
using the same format as for the SLR for the second expert report. The tables will be
automatically generated using the central database. Within this table the studies will
be ordered according to design (e.g. randomised controlled trials, cohort studies).
Within this table the studies will be ordered chronologically.

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: Energy intake
C: Smoking
D: Anthropometry: BMI, others

E: Physical activity
F: Alcohol
G: Others, e.g. occupational exposure, ethnicity, previous renal diseases, etc.

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.

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. When the number of studies justifies it, these results will be also presented in forest plots.

References

1. World Cancer Research Fund/ American Institute for Cancer Research: Food, Nutrition and the Prevention of Cancer: a global perspective. Washington DC: American Institute for Cancer Research, 1997
2. World Cancer Research Fund/ American Institute for Cancer Research: Food, Nutrition, Physical Activity, and the Prevention of Cancer: a global perspective. Washington DC: American Institute for Cancer Research, 2007
3. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. JAMA 1995; 274:1800-4
4. 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. Available from: http://globocan.iarc.fr, accessed on 13/10/2011]
5. Chow WH, Dong LM, Devesa SS. Epidemiology and risk factors for kidney cancer. Nat Rev Urol. 2010;7(5):245-57.
6. Ljungberg B, Campbell SC, Cho HY, Jacqmin D, Lee JE, Weikert S, Kiemeney LA. The epidemiology of renal cell carcinoma. Eur Urol. 2011;60(4):615-21.
7. Naito S, Tomita Y, Rha SY, Uemura H, Oya M, Song HZ, Zhong LH, Wahid MI. Kidney Cancer Working Group report. Jpn J Clin Oncol. 2010;40 Suppl 1:i51-56.
8. Levi F, Ferlay J, Galeone C, Lucchini F, Negri E, Boyle P, La Vecchia C. The changing pattern of kidney cancer incidence and mortality in Europe. BJU Int. 2008;101(8):949-58.
9. Mathew A, George PS, Ildaphonse G. Asian Pac J Cancer Prev. Obesity and kidney cancer risk in women: a meta-analysis (1992-2008). 2009; 10(3):471-8.
10. Ildaphonse G, George PS, Mathew A. Obesity and kidney cancer risk in men: a meta-analysis (1992-2008). Asian Pac J Cancer Prev. 2009;10(2):279-86
11. Lee JE, Männistö S, Spiegelman D, Hunter DJ, Bernstein L, van den Brandt PA, Buring JE, Cho E, English DR, Flood A, Freudenheim JL, Giles GG, Giovannucci E, Håkansson N, Horn-Ross PL, Jacobs EJ, Leitzmann MF, Marshall JR, McCullough ML,Miller AB, Rohan TE, Ross JA, Schatzkin A, Schouten LJ, Virtamo J, Wolk A, Zhang SM, Smith-Warner SA. Intakes of fruit, vegetables, and carotenoids and renal cell cancer risk: a pooled analysis of 13 prospective studies. Cancer Epidemiol Biomarkers Prev. 2009;18(6):1730-9
12. Faramawi MF, Johnson E, Fry MW, Sall M, Zhou Y. Consumption of different types of meat and the risk of renal cancer: meta-analysis of case-control studies. Cancer Causes Control. 2007;18(2):125-33.
13. Lee JE, Spiegelman D, Hunter DJ, Albanes D, Bernstein L, van den Brandt PA, Buring JE, Cho E, English DR, Freudenheim JL, Giles GG, Graham S, Horn-Ross PL, Håkansson N, Leitzmann MF, Männistö S,

McCullough ML, Miller AB, Parker AS, Rohan TE, Schatzkin A, Schouten LJ, Sweeney C, Willett WC, Wolk A, Zhang SM, Smith-Warner SA. Fat, protein, and meat consumption and renal cell cancer risk: a pooled analysis of 13 prospective studies. J Natl Cancer Inst. 2008;100(23):1695-706.

14. Lee JE, Hunter DJ, Spiegelman D, Adami HO, Bernstein L, van den Brandt PA, Buring JE, Cho E, English D, Folsom AR, Freudenheim JL, Gile GG, Giovannucci E, Horn-Ross PL, Leitzmann M, Marshall JR, Männistö S, McCullough ML, Miller AB, Parker AS, Pietinen P, Rodriguez C, Rohan TE, Schatzkin A, Schouten LJ, Willett WC, Wolk A, Zhang SM, Smith-Warner SA. Intakes of coffee, tea, milk, soda and juice and renal cell cancer in a pooled analysis of 13 prospective studies. Int J Cancer. 2007;121(10):2246-53.

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

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

17. N Orsini, R Bellocco and S Greenland, Generalized least squares for trend estimation of summarized dose-response data, Stata *J* 6 (2006), pp. 40–57

18. DerSimonian R, Laird N Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–188

19. Bekkering GE et al. How much of the data published in observational studies of the association between diet and prostate or kidney cancer is usable for meta-analysis? Am J Epidemiol 2008;167(9):1017-26.

20. Chêne G, Thompson SG. Methods for summarizing the risk associations of quantitative variables in epidemiologic studies in a consistent form.Am J Epidemiol. 1996;144(6):610-21

21. A Tobias. Assessing the influence of a single study in meta-analysis, Stata Tech Bull 47 (1999), pp. 15–17.

22. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. European Journal of Epidemiology 25 (9): 603-605, DOI: 10.1007/s10654-010-9491-z.

23. Cochrane Handbook for Systematic Reviews of Interventions [Internet] Available from http://www.cochrane.org/training/cochrane-handbook.

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

25. Annex 1.
WCRF - PUBMED SEARCH STRATEGY

a) 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 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]

#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]

#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

b) Searching for all studies relating to kidney cancer:

#23 renal neoplasma [MeSH terms] OR kidney neoplasma [MeSH terms] OR urethral neoplasma [MeSH terms]

24 kidney OR renal OR urethral AND (malign* [tiab] OR cancer*[tiab] OR carcinoma*[tiab] OR tumor*[tiab] OR tumour*[tiab] OR adenocarcinoma* OR sarcoma)

25 hypernephroma* OR grawitz tumo* OR wilms tumo* OR nephroma*

#26: # 23 OR #24 OR #25

c) Searching for all studies relating kidney cancer, and food, nutrition and physical activity:

#27 #22 AND #26

Annex 2. List of exposure codes (new sub-exposure codes indicated with *)

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

Measured in	Include	Exclude
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 in	Include	Exclude
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 sources (p 193)	Yes (p 197)
Alpha-carotene Beta-cryptoxanthin Lutein+zeaxanthin Lycopene	Plasma	Yes	Yes (p 194)	There is some evidence for interaction between carotenoids during intestinal absorption, which may complicate relationship between intake and blood levels (p 198)
Vit E	Plasma	Yes	Yes (p 196) NB. Strong confounding with serum cholesterol and total lipid concentrations (p 196).	Plasma, red and white blood cells. Yes, if used for vit E supplements. Yes, although if used for diet, associations are only moderate (p199)

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

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

Exposure	Measured in	Valid?	Reason (Willett)	Reason (Margetts / Nelson)
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	No	Plasma linoleic acid can discriminate between groups with relatively large differences in intake but performs less well on an individual basis (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
	Adipose tissue	Yes	Yes (p 220)	
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	
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

Exposure	Measured in	Valid?	Reason (Willett)	Reason (Margetts / Nelson)
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).

Nutrient	Biologic tissue	Val./reproduc	Coef	Details
Vitamin C	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)
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)

Nutrient	Biologic tissue	Val./reproduc	Coef	Details
Selenium	Whole blood	Validity	0.62	Correlation between selenium intake and whole blood selenium in South Dakotans (n=44) (Willett, p 186)
		Reproducibility	0.95	Average correlation between repeated measurements at four 3-month intervals in 78 adults (Willett, p 188)
Linoleic acid	Adipose tissue	Validity	0.57	Correlation between dietary linoleic acid intakes determined from 7-day weighted diet records and the relative proportion of linoleic acid in adipose tissue in Scottish men (n=164). Also correlation between linoleic acid measured in adipose tissue and calculated from FFQ in 118 Boston-area men (Willett, p 220)
Eicosapentaenoic (n-3)	Adipose tissue	Validity	0.40	Correlation with intake estimated from three 7-day weighted food records (Willett, p 223).
		Reproducibility	0.68	Correlation over 8 months in 27 men and women aged 20-29 (Willett, p 223).
	Plasma	Validity	0.23	Correlation of cholesterol ester fraction and intake in 3,570 adults (Willett, p 223)
		Reproducibility	0.38	Correlation of two measurements taken 6 years apart in study of 759 Finnish youths (Willett, p 219)
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