Study protocol – The consumption of ultra-processed foods and cardio-metabolic health – A systematic literature review.

GENERAL INFORMATION

Topic / title: Ultra-processed foods and cardio-metabolic health – A systematic literature review.
Registration: Will be registered at PROSPERO (currently no similar review registered).
Authors: Kai Schulze, Martin White, Jean Adams, Tarra Penney, Rebecca Love.
Financial support: MRC PhD studentship.

BACKGROUND RATIONALE

Population dietary patterns change as a result of economic development, increases in urbanization, and modernization, a process that is commonly referred to as the “nutrition transition”. Previous research has represented this transition in five different patterns, beginning with the hunter gatherer pattern to a pattern in which the results of overeating, diets high in processed and energy-dense foods, and obesity-related diseases (pattern four) are addressed by behaviour changes to prevent diet-related chronic health conditions (pattern five). Currently, most industrialized countries can be placed in pattern four with some elements of pattern five, whereas most low- and middle-income countries shift from the “end of famine” (pattern three) to the fourth pattern. Globally, the nutrition transition has indeed been paralleled by the epidemiological transition, with diet now constituting a leading cause of poor health. Dietary risks such as diets low in fruit, nuts, seeds or whole grains, and diets high in sodium, sugar-sweetened beverages, trans fatty acids or processed meats now account for the most DALYs (10%) of all risk factors, according to the Global Burden of Disease Study 2013.

At the same time, processed and ultra-processed foods now dominate the global food system. In Brazil, the United States, and Canada, 30%, 58%, and 61.7% of overall dietary energy and between 80% and 90% of all added sugars were consumed through ultra-processed foods. Similar figures can be found in Europe, where highly industrially processed foods account for between 53% (UK) to 75% (Germany, Netherlands) of the overall calorie intake. Foods with a higher degree of processing have been characterized by a disadvantageous nutritional profile, including a high energy-density combined with few nutrients per calorie (“empty calories”), a high content of trans-fatty acids from hydrogenated oils, sodium, (added) sugars, artificial food additives, a lack of fibre and phytochemicals, and a lack of micronutrients. The spread and domination of ultra-processed foods (UPF) has therefore been singled out as an indicator of the nutrition transition and a key contributor to the epidemic of obesity and other diet-related cardio-metabolic conditions.

On the observational individual level, a systematic epidemiological evidence base has been established for the associations between sugar-sweetened beverages (SSB) and obesity and diabetes. SSB’s will hence not be the focus of this review. Regarding UPF excluding SSB, however, the epidemiological evidence is less well defined. There are multiple reasons for this. First, there are many ways in which processed foods have been described in previous research. Fast foods, convenience foods, junk foods, highly (industrially) processed foods, snack foods, foods from a “western diet” or ultra-processed foods all refer to roughly the same type of foods (high in saturated / trans fats, ‘empty’ carbohydrates, not fruits and vegetables or whole grains). This multitude of concepts renders the building of a coherent evidence base challenging. A further complication has been added given the relatively recent academic interest in the degree of food processing and its relationship to health. Beginning in 2009, some authors including Carlos Monteiro and colleagues, criticised the dominating food group classifications for their difficulties in establishing convincing
associations between food groups and health outcomes.\textsuperscript{23,24} They claimed that food classifications based on processing are more relevant to nutrition and health when compared to classifications based on botanical or animal origin. Based on the hypothesis that diets high in ultra-processed foods should be linked to health, they proposed a new food classification system based on the degree of processing – the NOVA classification.\textsuperscript{24,25}

Since then, a limited number of studies have based the dietary exposure based on a systematic food processing classification and investigated the links to health. For example, positive associations between the consumption of UPF and obesity have been found in Brazil, UK, the US, Spain, and Sweden.\textsuperscript{11,9,26–31} Severe limitations for causal inference apply, however. Some of the studies rely purely on analyses on the ecological-level or anecdotal evidence. When trends in consumption of UPF and trends in health are found to have happened in parallel on the ecological level, an association is assumed or implied, although no formal modelling and taking into account of confounding factors has taken place. The studies that employ a comprehensive food classification based on food processing on the individual observational level are mostly cross-sectional and almost exclusively address the links to obesity but no other cardio-metabolic conditions. For certain types of dietary intake assessments, including some versions of food frequency questionnaires, the problem is that they have not been designed to comprehensively characterize dietary patterns according to the degree of food processing, which renders some studies or surveys unsuitable.\textsuperscript{24}

Yet, even though evidence regarding the specific food group termed “ultra-processed foods” and health is still inconclusive, relevant evidence that should be included in the debate around ultra-processed foods potentially does exist but is missed due to the focus on terminology. For example, a comprehensive study of three cohorts in the United States showed that the consumption of different food products such as desserts, fries, crisps, cookies or cakes, white bread, and sweets was associated with weight gain in adults.\textsuperscript{32} These foods are all ultra-processed, but the primary goal of the study was not to investigate the degree of food processing but to compare different dietary and lifestyle risk factors. Hence, it contained no reference to the degree of processing or ultra-processed foods in the title or abstract. Nonetheless, it did include valuable information about food groups that would be classified as highly or ultra-processed, which would be missed in a review that focused on the term ultra-processed foods solely.

Furthermore, much of the research in the field of nutrition has focused on single nutrients, foods, and/or food groups. While looking at components of the diet individually is important to examine the effects of various aspects of the diet on health, foods and nutrients are eaten in a variety of combinations and can have interactive and potentially cumulative or confounding relationships. Dietary pattern analysis facilitates the investigation of the effects of overall diet on health by attempting to characterise diets comprehensively. The underlying assumption of this review is that diets high in ultra-processed foods are adversely associated with cardio-metabolic health outcomes. Dietary patterns will therefore be the focus of the review, studies that only analyse single ultra-processed foods groups will not be considered.

Dietary patterns can be assessed in a number of ways, including numerical indices designed to gauge adherence to a particular pattern (e.g., Healthy Eating Index [HEI]) or data-driven approaches that use mathematics to empirically derive food intake patterns inherent among the study population (e.g., factor or cluster analysis). Dietary patterns can also be tested in trials or observed in observational studies. Because each methodology provides information about dietary patterns from a different perspective, the systematic review questions included in this project are based on the prevalent dietary pattern assessments in the literature: (1) index analysis, (2) factor/cluster analysis, (3) reduced rank regression.
OBJECTIVE AND RESEARCH QUESTION

To conduct a systematic literature review about the association between dietary patterns that are characterized by a higher intake of highly or ultra-processed food products and cardio-metabolic health.

Research questions:

Main research question:

What is the association between dietary patterns characterized by a higher share of highly or ultra-processed food consumption and cardio-metabolic health in human aged older than two years?

This research question is adapted to three different types of studies:

1. A priori index

What is the relationship between dietary patterns characterized by a higher share of ultra-processed foods assessed using an a priori defined index, score, classification or other numerical indices, and cardio-metabolic health outcomes?

2. A posteriori - Factor analysis, principal component analysis; cluster analysis

Are prevailing patterns of diets that are characterized by a higher share of ultra-processed foods related to cardio-metabolic health outcomes?

3. Reduced rank regression; discriminant analysis

Are there combinations of ultra-processed food intake in dietary patterns that explain the most variation in a risk of cardio-metabolic health outcomes?

METHODS

Study design

The systematic review will be conducted and reported in accordance with the PRISMA guidelines for the reporting of systematic reviews and meta-analysis. Table 1 shows the different stages that will be included in the review.

Table 1: Stages of the systematic review

<table>
<thead>
<tr>
<th>Stage</th>
<th>Authors involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Identification of records through database searching and additionally through other sources</td>
<td>KS</td>
</tr>
<tr>
<td>2. Screening based on abstract and title</td>
<td>KS+JA+ MW+TP+RL</td>
</tr>
<tr>
<td>3. Eligibility assessment of full-text articles and exclusion</td>
<td>KS+JA+ MW+TP+RL</td>
</tr>
<tr>
<td>4. Confirmation final studies for inclusion in qualitative and quantitative analysis</td>
<td>KS+JA+ MW+TP+RL</td>
</tr>
<tr>
<td>5. Data extraction and management</td>
<td>KS+JA+ MW+TP+EW</td>
</tr>
<tr>
<td>6. Quality and risk of bias assessment of included studies</td>
<td>KS+JA+ MW+TP+EW</td>
</tr>
<tr>
<td>7. Graphic and narrative synthesis</td>
<td>KS+JA+ MW+TP+EW</td>
</tr>
</tbody>
</table>
Databases

Several sources will be searched for published studies that are relevant to investigate the research question. All sources will be searched without restriction by location with the pre-piloted search strategy.

Consistent with the Cochrane Collaboration methodology for conducting a comprehensive search of the literature, the search strategy was developed to be both inclusive and rigorous and pre-piloted.\textsuperscript{34} To identify primary studies, searches in four databases that were suggested by an experienced librarian will be conducted:

- Ovid Medline
- Ovid Embase
- Cochrane library
- CINAHL

Study inclusion and exclusion criteria

The systematic review will include studies that estimate the association between the consumption of processed foods and cardio-metabolic health outcomes. No restriction will be placed on publication year, individual characteristics of study participants and geographic location.

The search will be restricted to peer-reviewed journal articles and quantitative study designs such as cohorts, cross-sectional or longitudinal, or trials. The focus will be on epidemiological studies; purely biological or chemical studies will not be considered (table 2).

Table 2: Inclusion and exclusion criteria

<table>
<thead>
<tr>
<th>Electronic search</th>
<th>Inclusion criteria</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
<td>Children, adolescents, and adults aged two years and older; of any gender, ethnicity or other characteristics; subjects free of outcome(s) of interest at onset of study</td>
<td>Animals and in vitro models, children &lt;2</td>
</tr>
<tr>
<td><strong>Exposure</strong></td>
<td>Intake of UPF in the context of overall dietary patterns (i.e. % of total energy, etc.)</td>
<td>SSB’s; UPF as single food group without the context of the overall dietary patterns</td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td>Objectively and subjectively measured cardio-metabolic health outcomes – obesity, overweight, T2DM, cardiovascular diseases, mortality related to those outcomes, and risk factors such as hypertension, raised triglycerides or low HDL-cholesterol</td>
<td>-</td>
</tr>
<tr>
<td><strong>Study Design</strong></td>
<td>Cohorts (cross-sectional, longitudinal), RCT’s, ecological studies</td>
<td>Systematic review, meta-analysis, narrative review, retrospective studies</td>
</tr>
<tr>
<td><strong>Publication type, year, and setting</strong></td>
<td>Peer reviewed journal articles of any year of any country</td>
<td>Grey literature, conference abstracts of unpublished studies</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td>No language restriction</td>
<td>-</td>
</tr>
<tr>
<td><strong>Additional</strong></td>
<td>Description of the dietary pattern(s) consumed by</td>
<td>-</td>
</tr>
</tbody>
</table>
Search strategy

The development of the search strategy has involved different steps. First, initial search terms that could potentially be used to describe the dietary exposure “ultra-processed foods” have been taken from a current systematic review about the classifications of the degree of food processing of foods. Further terms have been compiled by retrieving the terms of the ten most consumed ultra-processed foods or food groups (in terms of energy, based on the NOVA classification) from a nationally representative cross-sectional study in the US. Additionally, the 10 ultra-processed foods with the highest global sales (total kg, NOVA classification) according to the Euromonitor database have also been included. Secondly, the search terms for the three main cardio-metabolic health outcomes and their risk factors – obesity, type 2 diabetes mellitus, and cardiovascular diseases, etc. – have been compiled from 9 recent, widely cited systematic reviews and meta analyses. Some of the risk factors and “sub-diseases”, such as hypertension, are incorporated by including the MeSH terms for CVD. Thirdly, after initial development of the search strategy, an experienced medical librarian (Isla Kuhn, Clinical School University of Cambridge) was consulted to review the strategy and discuss the selection of databases. The strategy was then tested in the four selected databases, adjustments to the strategy were made and then again reviewed by the librarian. Table 3 shows the search terms and the results of the pre-piloted search strategy, and table 4 the example of the application of the search terms in OVID Medline.

Table 3: Pilot search terms (adapted for each database)

<table>
<thead>
<tr>
<th>Search terms 1 (dietary exposure – processed / ultra-processed foods):</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) (“processed” or “ultraprocessed” or “convenience” or “snack” or “fast” or “junk” or “packaged” or “chilled” or “frozen” or “refined” or “canned” or “industrial*” or (ready#prepared) or (ready#eat) or “ready-to-eat”)</td>
</tr>
<tr>
<td>b) (“food” or “foods” or “product” or “products” or “meal*” or “food product*” or “foodstuff*”)</td>
</tr>
<tr>
<td>c) Search: a) adjacent by maximum of four words with b)</td>
</tr>
<tr>
<td>d) (“snack*” or “ready#meal*” or “confectionary” or “bread*” or “baked good*” or “pizza*” or “icecream*” or “biscuit*” or “breakfast cereal*” or “cake*” or “cookie*” or “pie*”)</td>
</tr>
<tr>
<td>e) Search: c) or d)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Search terms 2 (association, relationship, link):</th>
</tr>
</thead>
<tbody>
<tr>
<td>(“relation*” or “related” or “association*” or “associated” or correlation*” or “correlated” or “connection*” or “connected” or “links” or “link” or “linked”)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Search terms 3 (obesity):</th>
</tr>
</thead>
<tbody>
<tr>
<td>(“obes*” or “overweight” or “overweight” or “BMI” or “body mass” or “body fat” or “body composition” or “body weight” or “body shape” or “waist circumference” or “abdominal fat” or “adiposity” or “waist circumference” or “skinfold” or “skin fold” or “waist to hip ratio” or “waist-hip ratio” or “waist to height ratio” or “waist-height ratio” or “weight adj2 gain” or “weight adj2 loss” or “weight adj2 loss” or “weight adj2 change”)</td>
</tr>
</tbody>
</table>

1 Note: MeSH or Emtree or “Exploding terms” are not included as they are not applicable to all databases. These terms are included in the search strategy for each database individually, see table 4 for the OVID Medline example.
Search terms 4 (diabetes):

a) (“diabetes”) AND (“type 2” or “type two” or “type ii” or “type II” or “non-insulin dependent”)
b) (“T2DM” or “prediabet*” or “pre-diabetes” or “prediabetic state” or “pre-diabetic” or “glucose intolerance” or “glucose intolerant” or “IGT” or “IFG”)
c) (“impaired fasting”) AND (“glucose” or “glycaemi*” or “glycemi*” or “bloodglucose” or “blood glucose”)
d) Search: a) or b) or c)

Search terms 5 (cardiovascular diseases):

a) (“cardiovascular disease*” or “CVD” or “coronary heart disease*” or “CHD” or “coronary disease*” or “cerebrovascular disease*” or “heart disease*” or “myocardial infarction” or “myocardial ischemia” or “acute coronary syndrome” or “stroke” or “haemorrhagic stroke” or “ischemic stroke” or “ischemic heart disease” or “ischaemic heart disease” or “blood pressure” or “hypertension” or “cholesterol” or “triglycerides”)

Search terms 6 (Exclusion of animal and in vitro studies):

a) (“animals” or “rat” or “rats” or “mouse” or “mice” or “animal study” or “animal studies” or “in vitro”)
b) (“human”) [i.e. MeSh terms for “human studies”]
c) Search: [a) not [a) and b)]

Search strategy:

a) (1 AND 2 AND (3 OR 4 OR 5)) NOT 6

Pilot Search Health Outcomes

<table>
<thead>
<tr>
<th>Database</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovid Medline</td>
<td>3856</td>
</tr>
<tr>
<td>Ovid Embase</td>
<td>5825</td>
</tr>
<tr>
<td>Cochrane Library</td>
<td>1703</td>
</tr>
<tr>
<td>CINAHL</td>
<td>945</td>
</tr>
<tr>
<td>Total</td>
<td>12329</td>
</tr>
<tr>
<td>De-duplicated</td>
<td>6011</td>
</tr>
</tbody>
</table>

Table 4: Example of Pilot Search in Ovid Medline

<table>
<thead>
<tr>
<th>#</th>
<th>A - Dietary exposure (processed food) search terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>processed.ti,ab.</td>
</tr>
<tr>
<td>2</td>
<td>ultraprocessed.ti,ab.</td>
</tr>
<tr>
<td>3</td>
<td>convenience.ti,ab.</td>
</tr>
<tr>
<td>4</td>
<td>snack.ti,ab.</td>
</tr>
<tr>
<td>A</td>
<td>Association search terms</td>
</tr>
<tr>
<td>---</td>
<td>--------------------------</td>
</tr>
<tr>
<td>1</td>
<td>exp Association/</td>
</tr>
<tr>
<td>2</td>
<td>relation/ ti,ab.</td>
</tr>
<tr>
<td>3</td>
<td>correlation/ ti,ab.</td>
</tr>
<tr>
<td>4</td>
<td>link ti,ab.</td>
</tr>
<tr>
<td>5</td>
<td>links/ ti,ab.</td>
</tr>
<tr>
<td>6</td>
<td>connection/ ti,ab.</td>
</tr>
<tr>
<td>7</td>
<td>(associated or correlated or linked or related or connected)/ ti,ab.</td>
</tr>
<tr>
<td>B</td>
<td>1 or 2 or 3 or 4 or 5 or 6</td>
</tr>
</tbody>
</table>
exp diabetes mellitus, type 2/
exp prediabetic state/
diabetes.ti,ab.
(Type 2 or type two or type ii or type II).ti,ab.
3 and 4
1 or 5
prediabetic state.ti,ab.
pre-diabetic.ti,ab.
7 or 8
D 6 or 9

# E- Cardiovascular disease search terms
1 exp cardiovascular diseases/
stroke.ti,ab.
(acute coronary syndrome).ti,ab.
exp coronary diseases/
(heart adj3 disease*).ti,ab.
(coronary adj3 disease*).ti,ab.
(cerebrovascular adj3 disease*).ti,ab.
exp heart diseases/
(CHD or CVD).ti,ab.
exp myocardial infarction/
exp myocardial ischemia/
E 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12

# F- Exclusion of animal and in vitro models
1 exp animals/
exp in vitro/
(rodent or rats or mice or mouse or animal study or animal studies or in vitro).ti,ab.
4 1 or 2 or 3
5 exp humans/
F (4 not (4 and 5))

Search strategy: (A AND B AND (C OR D OR E)) NOT F

Study selection – includes screening for eligibility

At each stage, independent dual screening will be conducted by the main reviewer (KS) and by another reviewer. Primary study titles identified in the initial search will be screened and those clearly outside the review criteria will be discarded. The abstracts of the remaining citations that pass the initial title screening will be reviewed and compared to the inclusion criteria to determine if retrieval of the full primary study is needed for further examination. A study will be included if it meets all the above inclusion criteria and excluded if it fails to meet all the inclusion criteria. If a study is not clear, it will be placed in the “maybe folder” and full text review will determine eligibility. At the full text review stage, any disagreements will be discussed between two reviewers and if consensus is not reached, a third reviewer will be approached. Each study included will have a record of decisions made at each stage.

Reference lists of included papers will be scanned for additional relevant articles. Subsequent citations of the additional articles will also be included. Additionally, if existing reviews on the same topic are discovered, they will be scanned for missing studies. At each stage of the review process, all data will be entered and managed using Mendeley reference manager. Each stage of the search and screening process will be reported using appropriate figures.

Quality assessment tool and addressing publication bias

Quality assessment will be conducted by the lead author and duplicated by an additional author. Inter-rater reliability will be determined and reported as a percentage of items without initial consensus. The quality assessment will be based on the National Institute of Health “Quality Assessment Tool for Observational Cohort and Cross-sectional Studies” and the principles outlined in Sanderson et al. (2007).44,45
Bias within the review can stem from publication bias, time lag bias, location bias, citation bias, language bias and outcome reporting bias. The best practices that are suggested by the OPEN consortium, which was established in 2011 to produce guidelines to minimise those issues, will be followed. Also, the Evidence Analysis Manual (2012) of the Academy of Nutrition and Dietetics will be used to assess selection, performance, detection, attrition and reporting bias. The Cochrane Collaboration’s tool for assessing risk of bias will be used for the assessment of RCT’s. Risk of bias will be assessed by two independent reviewers.

**Data extraction and management**

For each of the primary studies that meet the inclusion criteria, characteristics and covariates will be extracted. The extracted data will include:

1. Study characteristics including the title, authors, date of publication, study population (e.g. sample size, age, gender etc.), setting and study duration.
2. Description of exposure variables / subject matter of interest and the method for measuring dietary exposure (UPF), what classification has been used, how UPF have been operationalized.
3. Outcome measures (relevant model statistics) and corresponding measures of precision (standard deviations, standard errors or 95% confidence intervals), adjusted and unadjusted; the method and statistical analysis models.
4. Measurements of study quality/risk of bias.

The extraction process will be performed manually using a pre-established data extraction form and Microsoft Excel by two reviewers independently. A data extraction form will be customized for the research questions by reviewing a sample of studies and will be discussed between reviewers before finalization.

**Data Analysis & Synthesis**

The results of the systematic review will be synthesized using narrative methods and, if possible graphic methods. It is yet unclear whether the included studies will be homogeneous enough to undertake a meta-analysis of the results. The potential for the usage of meta-analysis methods such as forest plots might be limited. If feasible, an adapted use of harvest plots will be used to graphically synthesize, display and assimilate the findings. A sensitivity analysis will be conducted to determine if all appropriate studies were captured in the search.

**Dissemination plans**

Results will be disseminated via international conferences, and publications.

**REFERENCES**

3. Popkin BM, Adair LS, Ng SW. Global nutrition transition and the pandemic of obesity in


