Dietary patterns and gastric cancer risk: a systematic review and meta-analysis

P. Bertuccio1,2*, V. Rosato1,2, A. Andreano2, M. Ferraroni2, A. Decarli2,3, V. Edefonti2 & C. La Vecchia1,2

1Department of Epidemiology, Istituto di Ricerche Farmacologiche Mario Negri, IRCCS, Milan; 2Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milan; 3Unity of Medical Statistics, Biometry and Bioinformatics, Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy

Received 4 December 2012; revised 4 February 2013; accepted 4 February 2013

Background: Declines in gastric cancer (GC) incidence and mortality have been related to improvements in diet. It is therefore important to consider dietary patterns.

Design: We conducted a systematic review and meta-analysis of the literature through Medline and Embase databases.

Results: We identified 16 papers, of these 9 derived dietary patterns through an a posteriori method, 5 through a priori scores, and 2 used both approaches. Eight studies that used the a posteriori approach were considered for the meta-analysis. A favorable role on GC emerged for the ‘Prudent/healthy’, with an odds ratio (OR) of 0.75 [95% confidence interval (CI): 0.63–0.90], for the highest versus the lowest category. Similar results emerged for separate anatomical subtypes. An unfavorable role on GC emerged for the ‘Western/unhealthy’ dietary pattern, with an OR of 1.51 (95% CI: 1.21–1.89). This association was weaker for the distal/NOS (not otherwise specified) category (OR = 1.36) compared with the cardia GC (OR = 2.05). Among the a priori scores, the ORs ranged from 0.2 to 0.7 for the favorable and from 1.8 to 6.9 for the unfavorable ones.

Conclusion: There is a ~2-fold difference in GC risk between a ‘Prudent/healthy’ diet-rich in fruits and vegetables, and a ‘Western/unhealthy’ diet-rich in starchy foods, meat and fats.

Key words: dietary patterns, gastric cancer, meta-analysis, review, score

Introduction

Gastric cancer (GC) mortality showed steady declines over most recent decades worldwide, with however substantial geographic variations [1]. The lowest mortality rates were observed in most countries of northern and western Europe, in the USA and Canada, and in a few central American countries (i.e. Mexico, Cuba and Puerto Rico). The areas with persistent highest rates include the Russian Federation and other countries of central and eastern Europe, Japan and the Republic of Korea, and some countries of Latin America (i.e. Chile, Costa Rica, Colombia and Ecuador). Gastric cancer represents the fourth most common cancer and the second-leading cause of cancer death worldwide, with almost 1 million cases and over 700 000 deaths estimated in 2008 [2]. The downward trends, however, do not apply to the cardia subtype [3].

The declines in GC mortality rates are mainly attributed to a declined prevalence of H. pylori infection in more recent generations, but also to improved diet, including diet variety and food preservation. In particular, a diet high in fruits and vegetables and low in starchy and salty foods may have a protective role [4–8]. Most studies considered single dietary components, foods, food groups or nutrients. More recently, dietary patterns have been applied [9] in order to summarize and capture the overall dietary exposure and variations.

We conducted, therefore, a systematic review of the literature on the relation between dietary patterns and risk of GC and, when available, distal GC and cardia cancer separately. We then conducted a meta-analysis to provide quantitative estimates of the association.

Methods

Search strategy and data extraction

We carried out a search through PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) and Embase (http://www.embase.com) to identify all the original articles in English on the association between GC risk and dietary patterns published up to July 2012, based on the following string (stomach OR gastric) AND neoplasm[MeSH Terms] AND (“dietary pattern**” OR “eating pattern**” OR “food**” OR “dietary habit**” OR diet OR “dietary”) AND (“factor analysis” OR “principal component analysis” OR “cluster analysis” OR clustering OR...
"reduced rank regression" OR "Mediterranean diet" OR "diet diversity" OR "diet variety" OR quality OR index" OR indices OR scores), following the guidelines from the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group [10]. Three authors, PB, VR and AA, independently selected the articles, retrieved and assessed the potentially relevant ones and checked their reference lists of interest to identify additional relevant publications. Discrepancies in data extraction were resolved by involving a fourth author (VE).

We focused on studies in which combinations of multiple food- and nutrient-related dietary components (dietary patterns) were considered. These can be identified according to three main approaches: (i) a posteriori, such as principal component analysis (PCA), factor analysis (FA), principal component factor analysis (PCFA) and cluster analysis (CA); (ii) a priori and (iii) ‘reduced rank regression’, i.e. an integration of the a priori and the a posteriori approaches. No studies were excluded for weakness of design or data quality.

For each study, we extracted the following information: study design, geographical area and country, number and anatomic subtype of cases, number of participants for cohort studies and number of controls for case–control ones, age range, data collection period, dietary patterns identification method (i.e. PCFA, PCA, FA, CA or dietary score), characteristics of the dietary assessment method, name and composition of the identified patterns, corresponding measure of association [odds ratio (OR) or rate ratio/relative risk (RR)] and adjustment variables. A Japanese study [11] presented results both including and excluding events in the initial years of follow-up; in this case, we opted for the results based on all participants.

Detailed results were reported in Supplementary tables, available at Annals of Oncology online, according to the following criteria: (1) only statistically significant risks; (2) risks referring to exposure variables in categories, when corresponding risks were also reported for continuous variables and (3) risks with the maximum number of adjustment variables included in the models, when results from models including different adjustment variables were presented.

**statistical analysis**

We conducted a meta-analysis to provide quantitative summary estimates of the association between dietary patterns and GC risk. We restricted the analysis to the a posteriori dietary patterns, evaluating the risk in the highest versus the lowest categories of the most common patterns identified using PCFA (or PCA or FA). When more than one publication was on the same population or subpopulation, we included in the meta-analysis only the most recent and informative one. Since the labeling of the patterns is somewhat arbitrary and the dietary patterns are population-specific, we considered only those patterns sharing most foods with similar factor loadings. In detail, we identified two common dietary patterns: the first one, which we named ‘Prudent/healthy’, was characterized by high loadings on a core subset of foods given by fruits and vegetables and tended to add extra foods like fish, poultry or potatoes. The examined studies labeled it as ‘Vitamin-rich’ [12], ‘Vegetable and fruit’ [11], ‘Healthy’ [13,14], ‘Prudent’ [15,16], ‘Vitamins and fiber’ [17] and finally two dietary patterns pooled together, labeled as ‘fruits/vegetables’ and ‘fish/Vitamin C’, from the same study [18]. The second pattern, which we named ‘Western/unhealthy’, showed high loadings on a core subset of several types of meat and tended to include also bread, high-fat dairy foods, eggs or sweets. The studies under consideration labeled it as ‘Refined’ [12], ‘Meat’ [11], ‘Western’ [13–16], ‘Animal products’ [17] and ‘Meat/nitrite’ [18].

We included in separate meta-analyses both ORs and RRs of the so-called ‘Prudent/healthy’ and ‘Western/unhealthy’, and we refer to them as RR estimates from now on. We included RRs for both distal or GC not otherwise specified (‘distal/NOS’) and cardia GC, separately. Summary RRs were calculated using random-effects models considering both within-study and between-study variations [19] and were presented using forest plots. The area of the square is proportional to the inverse of the variance of the natural logarithm of the RR, thus giving a measure of the amount of information available from each estimate. A diamond is used to plot the summary RRs, the center of which represents the RR estimates and the extremes show the 95% confidence intervals (CIs).

Since some studies provided sex-specific risk estimates [14–16], we pooled those single estimates using fixed-effects models. One study from the USA [18] provided two patterns (i.e. ‘Fruits/vegetable’ and ‘Fish/vitamin C’) that are amenable to inclusion in the meta-analysis for the ‘Prudent/healthy’ pattern, and therefore, we pooled the two separate risk estimates to obtain an unique study-specific estimate for this pattern. One Canadian study [15] provided also stratified analyses by subtype, and, in this case, we included in the meta-analysis the separate cancer subtype-specific risk estimates.

We assessed heterogeneity among studies using the $\chi^2$ test (defining a heterogeneity significant with a P-value of <0.10) [20] and quantified the inconsistencies using the $I^2$ statistic, which represents the percentage of the total variation across studies that is attributable to heterogeneity rather than to chance [21]. We carried out sensitivity analyses conducting subgroup analyses by study design, geographic area (i.e. Europe, Asia and America), type of dietary items used to identify the dietary patterns (i.e. food groups or nutrients) and statistical modeling strategies (i.e. one or more dietary patterns in the same model). Publication bias was assessed through the visual inspection of the funnel plots and quantified by the Egger’ [22] and the Begg’ test [23]. Statistical analyses were conducted using STATA version 11.

**results**

**study selection**

From the literature search through PubMed and Embase databases, we identified and screened 684 articles. Of these, 24 were considered of interest and their full texts were retrieved for detailed evaluation. Eight papers were subsequently excluded because they met the exclusion criteria: two studies analyzed a subgroup of patients included in more recent publications, two investigated mortality instead of incidence,
one was a review summarizing available evidence from the EPIC cohort study and three mentioned the term pattern(s), but considered single foods and/or nutrients. Thus, a total of 16 studies were included in our systematic review, and of these, 8 were considered for the meta-analysis, i.e. the studies based on the a posteriori dietary patterns identified by PCFA, PCA or FA (Figure 1).

**systematic review**

Table 1 shows the main characteristics of the 16 studies considered [11–18,24–31]. The papers included 4 cohort studies [11,14,30,31] and 12 case–control ones [12,13,15–18,24–29]. Of these, six were population-based [12,13,15,18,27,29] and six hospital-based [16,17,24–26,28]. They were published between 1985 and 2012. Nine studies were conducted in Europe [12,13,17,24–26,29–31], three in North America [15,18,27], two in South America [16,28] and two in Asia [11,14]. Data from two studies [30,31] were based on the EPIC cohort. Two papers, an Italian case–control study published in 1987 [25] and an Uruguayan study published in 2004 [28], have been updated in more recent publications [16,26]. We included in the systematic review both the earlier and the updated papers as they provided different analyses.

Three cohort [11,14,30] and seven case–control [12,16,17,24–26,28] studies reported on the associations between dietary patterns and GC without specification of the topographic origin. One cohort [31] and one case–control [27] study reported on the association for distal GC only. Among the other case–control studies, one [13] reported the risk of cardia GC only, one [18] reported data on distal and cardia GC separately and other two [15,29] reported on the association with the risk of GC overall and that of distal and cardia, separately.

Nine of the identified studies derived dietary patterns through an a posteriori method [11–14,16–18,27,29], five reported associations for dietary scores defined a priori [24–26,30,31] and two used both approaches [15,28].

Among the studies that identified a posteriori dietary patterns, two did not find any significant association with GC risk, one using CA [27] and one using PCA [11]. Most of the other nine studies identified a significant association between GC risk and at least two dietary patterns, one [12,14–17] (or two [28]) favorable and one [12,14–16,28] (or two [17]) unfavorable patterns. The remaining three studies [13,18,29] identified a single unfavorable pattern. The significant favorable dietary patterns were mainly characterized by high consumption of fruits and vegetables [28], or nutrients contained in such foods [12,17], plus other foods, including fish [14–16]. They were named as ‘Healthy’ [14,28], ‘Prudent’ [15,16] and ‘Vitamin-rich/ﬁber’ [12,17]. Corresponding significant RRs for GC ranged from 0.4 to 0.6 (for further details, see Supplementary Table S1, available at Annals of Oncology online). The significant unfavorable dietary patterns were mainly characterized by a diet rich in starchy foods, red and processed meat, cheese and butter. They were named as ‘Traditional’ [12,14,16], ‘Western’ [13,15], ‘Starchy’ [28] or ‘Starch-rich’ [17] and ‘Meat/nitrite’ [18]. Corresponding significant RRs for GC ranged from 1.7 to 3.0 (for further details, see Supplementary Table S1, available at Annals of Oncology online).

Concerning the a priori patterns, one study evaluated the role of the ‘relative Mediterranean Diet’ (rMED) score [30], two were based on Diet Diversity scores [26,31] and four showed empirical dietary scores based on the results obtained from preliminary analyses on single food items [15,24,25,28]. We reported further details in the Supplementary Table S2, available at Annals of Oncology online. In brief, a significant

---

**Figure 1.** Systematic review flowchart.

---
Table 1. Main characteristics of the epidemiological studies on the association between gastric cancer and dietary patterns defined using the a priori and the *a posteriori* approaches

<table>
<thead>
<tr>
<th>Reference</th>
<th>Geographical area: country (data collection period)</th>
<th>Number of cases</th>
<th>Number of subject at risk/Number of controls</th>
<th>Age (years)*</th>
<th>Type of dietary patterns/identification method</th>
<th>Main results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cohort studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeurnink et al. (2012) [31]</td>
<td>Europe: 10 countries&lt;sup&gt;h&lt;/sup&gt; (recruitment: 1992–2000)</td>
<td>180 OGA</td>
<td>452 269</td>
<td>35–70</td>
<td>A priori</td>
<td>DSD&lt;sub&gt;vegfr&lt;/sub&gt;:↑</td>
</tr>
<tr>
<td><strong>Case-control studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Reference</th>
<th>Geographical area: country (data collection period)</th>
<th>Number of cases</th>
<th>Number of subject at risk/Number of controls</th>
<th>Age (years)*</th>
<th>Type of dietary patterns/dietary patterns identification method</th>
<th>Main results</th>
</tr>
</thead>
</table>

<sup>a</sup> Age of cases and controls.  
<sup>b</sup>Dietary pattern used in the meta-analysis as the ‘Prudent/healthy’ pattern.  
<sup>c</sup>Dietary pattern used in the meta-analysis as the ‘Western/unhealthy’ pattern.  
<sup>d</sup>The 10 European countries included in the EPIC study were: UK, France, Denmark, Sweden, Germany, Italy, Spain, the Netherlands, Norway and Greece.  
<sup>e</sup>The paper provided a main analysis based on 1169 cases of GC and 2332 controls, plus a stratified analysis by subsite on 132 cases of gastric cardia adenocarcinoma and 161 of non-cardia gastric adenocarcinoma, respectively, and 886 controls, as calculated from the Table on subsite analyses.  
<sup>f</sup>The paper provided a main analysis based on 591 cases of GC, plus a stratified analysis by subsite on 72 cases of gastric cardia adenocarcinoma and 444 of non-cardia gastric adenocarcinoma, and 1463 overall set of controls.  

CA, cluster analysis; DSSvegfr, Diet Diversity Scores for vegetable and fruit consumption; EPIC, European Prospective Investigation into Cancer and Nutrition; FA, factor analysis; GC, gastric cancer; GCA, gastric cardia adenocarcinoma; H, hospital-based; JACC, Japan Collaborative Cohort study for Evaluation of Cancer Risk; JPHC, Japan Public Health Center-based prospective study; NA, not available; OGA, non-cardia gastric adenocarcinoma; P, population-based; PCA, principal component analysis; PCFA, principal component factor analysis; rMED, relative Mediterranean diet.
protective role emerged for the 'Diversity score' reported by an Italian study (OR: 0.7) [26]. A study conducted in Uruguay [28] reported a significant inverse association between GC risk and the 'Protective score' (OR: 0.2). A Canadian study [15] showed a significant inverse association between the 'Food Index score' and the risk of GC among both women and men (ORs: 0.4 and 0.6, respectively), the risk of cardia GC among men (OR: 0.6) and that of distal GC among women (OR: 0.4). Moreover, a study from the EPIC cohort [30] reported that a high compared with low adherence to the rMED was associated with a significant reduction in GC risk (HR: 0.6). In contrast, significant unfavorable scores were found in a cohort and three case–control studies. A recent cohort study found a positive association with the risk of distal GC and the 'Diet Diversity score' for vegetable and fruit consumption (HR: 1.8, borderline significant) [31], though the risk for the continuous variable did not show a significant association. In a case–control study from Greece [24], a data-derived 'Risk score' was positively associated with GC risk (RR: 41.2). Similarly, in an Italian case–control study [25], a positive association with GC risk was found for the 'Combined score' (OR: 6.9). Finally, a Uruguayan case–control study [28] reported a significant positive association between GC risk and the 'Risk Enhancing score' (OR: 3.2).

meta-analysis
Figure 2 shows the associations between the highest intake level (tertile or quartile) and the lowest one of the 'Prudent/healthy' dietary pattern and GC risk for the eight studies included in the meta-analysis, separately by distal/NOS and cardia GC and overall combined. When all studies were combined, a negative association emerged between the 'Prudent/healthy' dietary pattern and GC risk, with a pooled OR of 0.75 (95% CI: 0.63–0.90), in the presence however of heterogeneity (P = 0.009, I² = 59%). This association was similar considering the risk of the distal/NOS category (OR = 0.75; 95% CI: 0.58–0.97) and that of cardia GC (OR = 0.76; 95% CI: 0.61–0.94).

Figure 3 shows the associations between the highest intake level (tertile or quartile) and the lowest one of the 'Western/unhealthy' dietary pattern and GC risk, separately by distal/NOS, and cardia GC and overall combined. When all studies were combined, a positive association emerged between the 'Western/unhealthy' dietary pattern and GC risk, with a pooled OR = 1.51 (95% CI: 1.21–1.89), with evidence of heterogeneity (P = 0.010, I² = 59%). This association was weaker for the risk of the distal/NOS category (OR = 1.36; 95% CI: 1.07–1.73) compared with that of cardia GC (OR = 2.05; 95% CI: 1.51–2.78).

publication bias
The funnel plot for the 'Western/unhealthy' pattern gave evidence of asymmetry (not shown), as confirmed by the corresponding statistical tests (Egger’s test, P = 0.014; Begg’s test, P = 0.060). There was no evidence of asymmetry for the 'Prudent/healthy' pattern (Egger’s test, P = 0.91; Begg’s test, P = 0.66). However, the small number of studies does not allow for a robust interpretation of these results.

sensitivity analyses
The two cohort studies included in the meta-analysis were the only ones from Asia, and both did not provide risk estimates.
for cardia GC [11,14]. Their pooled estimates were not significant (RR: 1.02 for the ‘Western/unhealthy’ pattern and 0.93 for the ‘Prudent/healthy’ one). When we analyzed results from the six case–control studies [12,13,15–18], the overall RRs were similar to those of the main analysis including the cohort studies, being 1.67 (95% CI: 1.32–2.10) for the ‘Western/unhealthy’ pattern and 0.71 (95% CI 0.58–0.88) for the ‘Prudent/healthy’ one. When we analyzed studies by the geographic area, we found similar positive associations among the European (RR: 1.62; 95% CI: 1.14–2.31) and the American (RR: 1.72; 95% CI: 1.22–2.43) studies for the ‘Western/unhealthy’ pattern, but an appreciable stronger negative association among the European (overall RR: 0.57; 95% CI: 0.46–0.70) than among the American (RR: 0.80; 95% CI: 0.63–1.02) studies for the ‘Prudent/healthy’ pattern. When we removed the two studies based on nutrient intakes [12,17], the overall RRs were 1.51 (95% CI: 1.16–1.96) for the ‘Western/unhealthy’ pattern and 0.82 (95% CI: 0.70–0.97) for the ‘Prudent/healthy’ one. Finally, when we analyzed studies considering composite models [16–18], the overall RRs were 1.70 (95% CI: 1.11–2.59) for the ‘Western/unhealthy’ pattern and 0.77 (95% CI: 0.55–1.07) for the ‘Prudent/healthy’ one.

discussion

This is to our knowledge the first systematic review on dietary patterns and GC. Among the nine studies based on a posteriori dietary patterns, the RRs for overall GC ranged from 0.2 to 0.7 for favorable ones and from 1.8 to 6.9 for the unfavorable ones.

Most of the evidence in our systematic review and meta-analysis is based on case–control studies that have some typical limitations of retrospective studies.

Most of the a posteriori dietary patterns considered in this review were identified through multivariate statistical methods (i.e. FA, PCFA and CA). These methods imply subjective aspects and can lead to different choices and consequently to different results. This makes it difficult to compare results across studies. For the meta-analysis, to minimize the risk of bias, first we concentrated only on studies applying FA. Second, we identified only those dietary patterns as similar as possible across studies, in terms of factor loadings of the dietary items considered. Such an approach was used by other researchers in similar analysis [32,33]. All the dietary patterns were derived from the Food Frequency Questionnaire (FFQ), but only some of these were tested for reproducibility [14,16,17,28] and validity [14,17,29]. Further, the list of dietary items investigated in the FFQs varied widely across studies, as well as the number of items considered for the exploratory analyses. Consequently, the dietary patterns could be influenced by different dietary habits across populations. However, for the a posteriori patterns, the similarity in the types of foods considered in various studies is of particular interest—apart from the terminology used in the definition of the favorable/unfavorable patterns.

With reference to confounding, whenever possible, we included in our meta-analysis multivariate RRs adjusted for all the major recognized risk factors for GC, i.e. age, sex, education, body mass index, smoking, family history, alcohol
drinking and total energy intake. This reassures against a relevant role of bias or confounding.

*H. pylori* infection is a necessary but not sufficient cause for distal GC [4,8,34–37]. Nevertheless, retrospective studies have a limited ability to analyze *H. pylori*. In fact, subjects developing gastric atrophy as a consequence of *H. pylori* infection and, who are at a greater risk of developing GC, may have a low antibody titer at cancer diagnosis [4,35, 6]. Even assuming that *H. pylori* infection is present in all cases and a proportion of controls, a proportion of cases would be preventable by appropriate diets and associated factors [4,38]. Only a few studies examined the potential modifying effects and interactions of *H. pylori* infection on diet-GC risk [39,40].

To assess the potential sources of the heterogeneity across studies, we carried out a sensitivity analysis, which showed consistent results among studies according to selected study characteristics. Further, most findings emerging from several a priori patterns were consistent with those from the *a posteriori* ones, and they are in broad agreement with most data on single dietary components [41–44].

In short, we pooled information from eight studies that identified two *a posteriori* dietary patterns in terms of single food or nutrient items mainly correlated to them. The dietary pattern, named ‘Prudent/healthy’, mainly based on high consumption of fruits and vegetables, may decrease by 25% the risk of GC. Similar results emerged by anatomical subtype. In contrast, a dietary pattern named ‘Western/unhealthy’, characterized by consumption of several types of meat, high-fat dairy foods, starchy foods and sweets, increased by ~50% the risk of GC. There is therefore a ~2-fold difference in GC risk between ‘Prudent/healthy’ and ‘Western/unhealthy’ dietary patterns. These differences underlines perspectives for GC prevention, particularly in populations with high prevalence of *H. pylori*, and consequently with high incidence and mortality for GC [1–4,6,8].

acknowledgements

The authors thank Mrs I. Garimoldi for editorial assistance.

funding

This work was supported by the Italian Ministry of Education (PRIN 2009 X8YCBN), the Italian Association for Research on Cancer (AIRC No. IG10068 and IG10415).

disclosure

The authors have declared no conflicts of interest.

references


