Coffee and cancer risk: A meta-analysis of prospective observational studies

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1038/srep33711</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:29407556">http://nrs.harvard.edu/urn-3:HUL.InstRepos:29407556</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Coffee and cancer risk: A meta-analysis of prospective observational studies

Anqiang Wang1,*, Shanshan Wang1,*, Chengpei Zhu2,*, Hanchun Huang1, Liangcai Wu1, Xueshuai Wan1, Xiaobo Yang1, Haohai Zhang1, Ruoyu Miao3, Lian He1, Xinting Sang1 & Haitao Zhao1

Meta-analyses on coffee and cancer incidence mainly restricted to limited cancers. We carried out a more comprehensive meta-analysis of cohort studies to explore association between coffee and most cancer types. We conducted comprehensive search and summarized relative risk (RR) and 95% confidence intervals for the highest versus lowest coffee intake and cancer using STATA12. We conducted dose-analysis if result suggested significant association. The publication bias was evaluated with begg’s and egger’s test. Finally, 105 individual prospective studies were included. Inverse associations were observed on oral, pharyngeal, colon, liver, prostate, endometrial cancer and melanoma, with RR 0.69 (95% CI = 0.48–0.99, I² = 73.4%, P = 0.044), 0.87 (95% CI = 0.78–0.96, I² = 28.4%, P = 0.007), 0.46 (95% CI = 0.37–0.57, I² = 0%, P = 0), 0.89 (95% CI = 0.84–0.93, I² = 30.3%, P = 0.003), 0.73 (95% CI = 0.67–0.80, I² = 0%, P = 0) and 0.89 (95% CI = 0.80–0.99, I² = 0%, P = 0.031) respectively. However, the relative risk for lung cancer is 2.18 (95% CI = 1.26–3.75, I² = 63.3%, P = 0.005). The summary relative risk for increment of 2 cups of coffee were RR = 0.73, 95% CI = 0.67–0.79 for liver cancer, RR = 0.97, 95% CI = 0.96–0.98 for prostate cancer and RR = 0.88, 95% CI = 0.85–0.92 for endometrial cancer. Accordingly, coffee intake was associated with reduced risk of oral, pharynx, liver, colon, prostate, endometrial cancer and melanoma and increased lung cancer risk.

Cancer is a major cause of morbidity and mortality worldwide. In 20121,2, there were 14.1 million new cancer cases and 8.2 million cancer deaths globally and the global burden from cancer has become significantly high. Although the strategy for cancer treatment has improved much, cancer is still the most difficult disease to overcome. Early diagnosis and treatment may provide good chance for cancer patients to improve overall prognosis. However, there is no promising survival for those who suffer from advanced cancer. Therefore, it is important to determine methods to prevent tumor occurrence. Nutrition, such as the consumption of citrus fruits, vegetables, and coffee, has also been considered to play an important role in reducing cancer risk3–5.

Coffee is one of the most popular beverages worldwide, and it has been speculated to decrease the risk of many types of cancers. Coffee is a complex mixture of many biologically active components, some of which may have anti-tumor effects. They include caffeine, cafestol, kahweol, and chlorogenic acid. Although many standardized meta-analyses5–18 between coffee intake and various cancers have been conducted, most of analyses restricted to limited types of cancers. In addition, there is no relevant meta-analysis to explore relationships between coffee intake and some types of cancer including melanoma and lymphoma. Furthermore, dose-response meta-analyses were only performed for liver cancer and prostate cancer19,20, we endeavor to conduct such analysis for cancer as many as possible. Besides, most meta-analyses included prospective studies and case-control studies, which may lead to the unconvincing results and conclusions. However, we included prospective studies into our study. In order to present the relatively obvious associations between coffee intake and the incidence of most of cancer types, we carried out a more comprehensive systematic updated meta-analysis of cohort studies to explore the

1Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. 2 Department of General Surgery, Tianjin Third Central Hospital, Tianjin Institute of Hepatobiliary Disease, Tianjin Key Laboratory of Artificial Cell, Tianjin, China. 3 Liver Center and The Transplant Institute, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to X.S. (email: sangxt@pumch.cn) or H.Z. (email: ZhaoHT@pumch.cn)
association between coffee consumption and most types of cancers. We try to provide a landscape of coffee and cancer incidence.

**Methods**

**Search strategy.** We conducted a computerized search of the literature on coffee and cancer from inception to July 2015. The search terms were (cancer OR tumor OR carcinoma) AND (coffee OR caffeine OR beverages OR diet OR drinking). Three major electronic databases (PubMed, EMBASE, and The Cochrane Library) were used to search the relevant literature without language restriction. Then, we classified these studies into different groups according to cancer type and conducted repeated searches. Finally, we obtained a relatively complete literature record for every type of cancer. Moreover, we reviewed the references from retrieved articles for additional studies (Fig. 1).

**Study selection.** The included studies had to be prospective cohort, case-cohort, or nested case control studies and had to contain the association between coffee intake and cancer incidence. Case control studies and cross-sectional studies were excluded. Studies concerning the coffee intake and the mortality of cancer were also excluded. We also excluded the analysis about cancer if there was no enough to conduct meta-analysis or no updated prospective studies for previous meta-analysis. Each study had to provide estimates of hazard ratio (HR) or relative risk (RR) with 95% confidence intervals or numbers of cases and controls to calculate those values. For cancers for which a dose-response analysis could be conducted, the quantitative measure of intake and the total number of cases and person years or relative information to calculate them were necessary. Two independent reviewers evaluated the eligibility of every study in a standardized manner. We resolved disagreements by consensus.

**Data abstraction.** We extracted data from each study. These data included the first author’s name, year of publication, country of origin, number of cases and participants, sex, age, types of cancer, duration of follow-up and comparison of exposure level and adjusted confounding variables. The relative risks or hazard ratio and 95% confidence intervals for the highest versus the lowest intake were also extracted.
Statistical analysis. We summarized the overall relative risk and 95% confidence intervals for the highest versus lowest intake of coffee and various cancers using the fixed or random effects models of STATA12 (StataCorp, College Station, TX, USA). The command is \texttt{metan logr logl loguci, effect(RR)}. Meta\-reg analysis was conducted when there were no or low heterogeneity and medium or high heterogeneity, respectively. For studies that reported results by cancer type or sex, we combined them using fixed effects models. For studies that provided no relative risks and 95% confidence intervals, we calculated these values based on the number of cases and controls.

We conducted dose-response analyses for coffee and some types of cancers rather than all tumor types. When the summary RR indicates a significant association between coffee intake and cancer risk, we would conduct dose-response analysis for these types of cancers. Furthermore, the included studies for dose-response analysis had to be at least three. According to a previously described method\textsuperscript{21,22}, information was required, such as the distribution of cases and person-years, variance of exposure levels and relative risks for at least three quantitative categories. The number of person-years was approximated from follow-up duration and number of subjects if the studies did not provide these values. We assigned the mid-point of the corresponding range of coffee intake as the exposure value in each category. When the highest category was open-ended, we assumed the width of the interval to be the same as the adjacent interval. When the lowest category was open-ended, we assigned zero as the lower boundary\textsuperscript{23}. If coffee intake was reported as weeks or months instead of days, we computed the corresponding days. If coffee intake was not measured in cups, we used 150 ml\textsuperscript{24} as a cup to recalculate the intakes to a common scale. We present the dose-response results for a two-cup daily increment of coffee consumption. Using restricted cubic splines with four knots at percentiles 5%, 35%, 65% and 95% of the distribution; we evaluated the curve linear relation between coffee intake and risks of liver cancer, prostate cancer and endometrial cancer. The P value was calculated to test the null hypothesis that the coefficients of the second and third splines are equal to zero\textsuperscript{24,25}. According to the result, we conducted linear or non-linear dose-response analysis.

We used the I² value to evaluate if heterogeneity could be explained by study differences rather than by chance\textsuperscript{26}. I²-values of approximately 25%, 50%, and 75% indicate low, moderate and high heterogeneity. We did not use scores to assess the quality of the studies. Instead, we conducted subgroup analyses to evaluate the impact on overall RRs by cancer type, sex, follow-up period and some adjusted confounders, such as alcohol and BMI. Moreover, when the number of included studies was more than 10 and a substantial heterogeneity was observed, we would conduct subgroup analysis and meta-regression analysis to explore the potential source of heterogeneity. Furthermore, when observed heterogeneity was more than 10 and a substantial heterogeneity was observed, we would conduct subgroup analysis and meta-regression analysis to explore the potential source of heterogeneity.

We evaluated publication bias using Begg’s\textsuperscript{27} and Egger’s\textsuperscript{28} tests, with obvious publication bias suspected when P < 0.10. And we conducted a Trim and fill analysis\textsuperscript{29} to assess the stability of overall relative risk when the results suggested obvious publication bias. We conducted all of the statistical analyses using Stata Statistical Software, version 12.0.

Results

Digestive system cancer. Oral, pharynx cancer. Highest versus lowest intake. Six cohort studies\textsuperscript{30–35} were included in the analysis (1395309 samples) of the highest versus lowest intake of coffee and oral, pharynx cancer. The study characteristics are presented (Stable 1a). The summary RR was 0.69 (95% CI = 0.48–0.99, P = 0.044) with high heterogeneity (I² = 73.4%, P = 0.002) (Fig. 2A). The results suggest no publication bias, with P = 1 for Begg’s test and P = 0.98 for Egger’s test. The subgroup analysis indicated that the inverse association was observed between coffee consumption and oral, pharynx cancer incidence in Asia and follow-up (>10 years) subgroup. The same relationship was observed in smoking adjustment subgroup, physical activity adjustment subgroup and total energy adjustment subgroup (Stable 1b).

Esophageal cancer. Highest versus lowest intake. Six cohort studies\textsuperscript{30,32–34,36,37} were included in the analysis (1395309 samples) of the highest versus lowest intake of coffee and esophageal cancer. The study characteristics are presented (Stable 1a). The summary RR was 0.86 (95% CI = 0.71–1.04, P = 0.124) with no heterogeneity (I² = 0%, P = 0.64) (Fig. 3A). The results suggest no publication bias, with P = 1 for Begg’s test and P = 0.69 for Egger’s test. The subgroup analysis suggested that no significant association was observed between coffee intake and esophageal cancer incidence in each subgroup (Stable 1b).

Stomach cancer. Highest versus lowest intake. Twelve cohort studies\textsuperscript{30,31,33,38–46} were included in the analysis (1305447 samples) of the highest versus lowest intake of coffee and stomach cancer. The study characteristics are presented (Stable 1a). The summary RR was 1.15 (95% CI = 0.96–1.37, P = 0.121) with medium heterogeneity (I² = 49.2%, P = 0.027) (Fig. 3B). The results suggest no publication bias, with P = 0.63 for Begg’s test and P = 0.85 for Egger’s test. The subgroup analysis indicated that coffee intake was associated with increased stomach cancer in USA and Asia. The same relationship was also observed in small number cases subgroup (<500 cases) and short follow-up subgroup (<10 years). The meta-regression analysis found no obvious to explain source of heterogeneity (Stable 1b).

Colorectal cancer. Highest versus lowest intake. Twenty-one cohort studies\textsuperscript{30,31,38,43,47–49} were included in the analysis (2141185 samples) of the highest versus lowest intake of coffee and colorectal cancer. The study characteristics are presented (Stable 1a). The summary RR was 0.96 (95% CI = 0.91–1.02, P = 0.175) with low heterogeneity (I² = 23.6%, P = 0.160) (Fig. 4A). The summary RRs for colon and rectal cancer were 0.87 (95% CI = 0.78–0.96, P = 0.007) and 0.94 (95% CI = 0.85–1.04, P = 0.236) (Fig. 2B). The results suggest no publication...
bias, with $P = 0.70$ for Begg's test and $P = 0.82$ for Egger's test. The subgroup analysis indicated that no significant association was observed between coffee intake and the risk of colorectal cancer in each subgroup. No substantial source of heterogeneity was found by meta-regression analysis (Stable 1b).

Pancreatic cancer. Highest versus lowest intake. Fifteen cohort studies were included in the analysis (1219019 samples) of the highest versus lowest intake of coffee and pancreatic cancer. The study characteristics are presented (Stable 1a). The summary RR was 1.02 (95% CI = 0.87–1.18, $P = 0.832$) with low heterogeneity ($I^2 = 16.2\%$, $P = 0.27$) (Fig. 4B). The results suggest no publication bias, with $P = 0.43$ for Begg's test and $P = 0.75$ for Egger's test. There was no significant association between coffee intake and the risk of pancreatic cancer in each subgroup through subgroup analysis. Meta-regression analysis found no factors explaining the source of heterogeneity (Stable 1b).
Liver cancer. Highest versus lowest intake. Nine cohort studies23,75–82 were included in the analysis (968517 samples) of the highest versus lowest intake of coffee and liver cancer. The study characteristics are presented (Stable 1a). The summary RR was 0.46 (95% CI = 0.37–0.57, P = 0) with no heterogeneity (I² = 0%, P = 0.44) (Fig. 5A). The results suggest no publication bias, with P = 0.18 for Begg’s test and P = 0.23 for Egger’s test. The inverse association was observed between coffee consumption and liver cancer incidence in most subgroups through subgroup analysis. However, there was no significant relationship between coffee intake and the risk of liver cancer in women subgroup (Stable 1b).

Dose-response analysis. We included seven studies23,75–80,82 for the dose-response analysis. There was no non-linear association between coffee intake and risk of liver cancer (P = 0.41 for non-linearity, Fig. 5A). A statistically significant inverse association was observed for liver cancer (RR = 0.73, 95% CI = 0.67–0.79) with an increased intake of 2 cups of coffee per day.

Tumors of the urinary system. Renal cancer. Highest versus lowest intake. Five cohort studies30,31,43,83,84 were included in the analysis (1036465 samples) of the highest group (samples/cases: 125065/144) versus lowest group (samples/cases: 232706/299) of coffee intake and renal cancer. The study characteristics are presented (Stable 2a). The summary RR was 0.79 (95% CI = 0.54–1.15, P = 0.226) with medium heterogeneity (I² = 49.8%, P = 0.09) (Fig. 6A). The results suggest no publication bias, with P = 0.22 for Begg’s test and P = 0.14 for Egger’s test. There was no significant association between coffee intake and renal cancer risk in each subgroup through subgroup analysis (Stable 2b). We did not found factors to explain the source of heterogeneity through subgroup analysis.

Bladder cancer. Highest versus lowest intake. Ten cohort studies30,31,38,85–91 were included in the analysis (340544 samples) of the highest versus lowest intake of coffee and bladder cancer. The study characteristics are presented (Stable 2a). The summary RR was 1.12 (95% CI = 0.94–1.34, P = 0.192) with medium heterogeneity (I² = 39.6%, P = 0.094) (Fig. 6B). The results suggest publication bias, with P = 0.016 for Begg’s test and P = 0.01 for Egger’s test. The summary RR was 1.04 (95% CI = 0.88–1.23, P = 0.669) after Trim and fill analysis. The subgroup analysis indicated that coffee intake was associated with increased bladder cancer in USA (Stable 2b). No factor could explain the source of heterogeneity through subgroup analysis and meta-regression analysis (Stable 2b).

Prostate cancer. Highest versus lowest intake. Fourteen cohort studies30,31,38,92–101 were included in the analysis (864012 samples) of the highest versus lowest intake of coffee and prostate cancer. The study characteristics are presented (Stable 2a). The summary RR was 0.89 (95% CI = 0.84–0.93, P = 0.003) with medium heterogeneity (I² = 30.3%, P = 0.14) (Fig. 5B). The results suggest no publication bias, with P = 0.83 for Begg’s test and P = 0.84 for Egger’s test. Through subgroup analysis, we found the inverse association between coffee intake and prostate cancer incidence in most subgroups. However, there was no significant relationship between coffee intake and the risk of prostate cancer in Canada and Asia (Stable 2b). The same associations were observed in follow-up (< 10 years) subgroup and some adjustment confounders including no physical activity adjustment subgroup and no BMI adjustment subgroup (Stable 2b). Meta-regression found no substantial factor to explain the source of heterogeneity.

Dose-response analysis. We included ten studies30,31,92,96–101 in the dose-response analysis. There was no non-linear association between coffee intake and risk of prostate cancer (P = 0.15 for non-linearity, Fig. 5B). A statistically significant inverse association was observed for prostate cancer (RR = 0.97, 95% CI = 0.96–0.98) with an increased intake of 2 cups of coffee per day.
Figure 5. Dose-response analyses between coffee intake and risk of liver cancer, prostate cancer and endometrial cancer. Relative risks of liver cancer (A), prostate cancer (B) and endometrial cancer (C) associated with coffee intake. Squares represent study-specific relative risk estimates (size of the square reflects the study-specific statistical weight, that is, the inverse of the variance); horizontal lines represent 95% CIs; diamonds represent summary relative risk estimates with corresponding 95% CIs.

Figure 6. Meta-analyses between coffee intake and risk of renal cancer and bladder cancer. Relative risks of renal cancer (A) and bladder cancer (B) associated with coffee intake. Squares represent study-specific relative risk estimates (size of the square reflects the study-specific statistical weight, that is, the inverse of the variance); horizontal lines represent 95% CIs; diamonds represent summary relative risk estimates with corresponding 95% CIs.
Female genital system neoplasm. Breast cancer. Highest versus lowest intake. Seventeen cohort studies\(^{30,31,43,102–115}\) were included in the analysis (997482 samples) of the highest versus lowest intake of coffee and breast cancer. The study characteristics are presented (Stable 3a). The summary RR was 0.99 (95% CI = 0.94–1.04, \(P = 0.619\)) with no heterogeneity (\(I^2 = 0\), \(P = 0.55\)) (Fig. 7A). The results suggest no publication bias, with \(P = 0.84\) for Begg’s test and \(P = 0.75\) for Egger’s test. No significant relationship was observed between coffee consumption and the risk of breast cancer in all subgroups through subgroup analysis (Stable 3b).

Ovarian cancer. Highest versus lowest intake. Nine cohort studies\(^{30,31,43,116–121}\) were included in the analysis (687017 samples) of the highest versus lowest intake of coffee and ovarian cancer. The study characteristics are presented (Stable 3a). The summary RR was 1.04 (95% CI = 0.90–1.20, \(P = 0.582\)) with low heterogeneity (\(I^2 = 23.7\%), \(P = 0.23\)) (Fig. 7B). The results suggest publication bias, with \(P = 0.02\) for Begg’s test and \(P = 0.009\) for Egger’s test. The summary RR was 0.96 (95% CI = 0.84–1.09) after Trim and fill analysis. The subgroup analysis indicated that there was no significant association between coffee intake and ovarian cancer risk in each subgroup. No factor could explain the source of heterogeneity through subgroup analysis (Stable 3b).

Endometrial cancer. Highest versus lowest intake. Twelve cohort studies\(^{30,31,43,116–121}\) were included in the analysis (1114002 samples) of the highest versus lowest intake of coffee and endometrial cancer. The study characteristics are presented (Stable 3a). The summary RR was 0.73 (95% CI = 0.67–0.80, \(P = 0\)) with no heterogeneity (\(I^2 = 0\), \(P = 0.58\)) (Fig. 5C). The results suggest no publication bias, with \(P = 0.19\) for Begg’s test and \(P = 0.16\) for Egger’s test. Subgroup analysis indicated that inverse significant association was observed between coffee consumption and endometrial cancer incidence in most subgroups. However, there was no significant relationship between coffee intake and the risk of endometrial cancer in no BMI adjustment subgroup (Stable 3b).

Dose-response analysis. We included eleven studies\(^{30,31,43,122–129}\) for the dose-response analysis. There was no non-linear association between coffee intake and risk of endometrial cancer (\(P = 0.69\) for non-linearity, Fig. 5C). A statistically significant inverse association was observed for endometrial cancer (RR = 0.88, 95% CI = 0.85–0.92) with an increased intake of 2 cups of coffee per day.

Other cancers. Lung cancer. Highest versus lowest intake. Four cohort studies\(^{30,31,38,131}\) were included in the analysis (103137 samples) of the highest versus lowest intake of coffee and lung cancer. The study characteristics are presented (Stable 4a). The summary RR was 2.19 (95% CI = 1.26–3.75, \(P = 0.005\)) with high heterogeneity (\(I^2 = 63.3\%), \(P = 0.04\)) (Fig. 2C). The results suggest no publication bias, with \(P = 0.73\) for Begg’s test and \(P = 0.39\) for Egger’s test.

Melanoma. Highest versus lowest intake. Six cohort studies\(^{30,31,43,132–134}\) were included in the analysis (773536 samples) of the highest versus lowest intake of coffee and melanoma. The study characteristics are presented (Stable 4a). The summary RR was 0.89 (95% CI = 0.80–0.99, \(P = 0.031\)) with no heterogeneity (\(I^2 = 0\), \(P = 0.42\)) (Fig. 2D). The results suggest no publication bias, with \(P = 0.71\) for Begg’s test and \(P = 0.40\) for Egger’s test. There was no significant association between coffee intake and melanoma risk through subgroup analysis. The subgroup analysis indicated no factor to explain the source of heterogeneity (Stable 4b).

Lymphoma. Highest versus lowest intake. Three cohort studies\(^{30,31,43}\) were included in the analysis (89897 samples) of the highest group (samples/cases: 32783/63) versus lowest group (samples/cases: 17229/22) coffee intake and lymphoma. The study characteristics are presented (Stable 4a). The summary RR was 1.23 (95% CI = 0.75–2.03, \(P = 0.415\)) with no heterogeneity (\(I^2 = 0\), \(P = 0.769\)) (Fig. 8). The results suggest no publication bias, with \(P = 1\) for Begg’s test and \(P = 0.18\) for Egger’s test.
Subgroup, sensitivity, and meta-regression analyses. We conducted subgroup analyses and meta-regression analyses based on sex, duration of follow-up, geographical location and adjusted confounders. Coffee intake was associated with a decreased risk of liver cancer in most subgroups, with no substantial heterogeneity between subgroups. However, the protective effect was absent in the female population (Stable 1b). Inverse association was observed for oral, pharynx cancer in Asia rather than Europe and USA (Stable 1b). Inverse associations were also observed for prostate cancer and endometrial cancer, with no substantial heterogeneity between subgroups (Stable 2b and 3b). Coffee intake was not significantly associated with colorectal cancer incidence in most subgroup analyses. However, an inverse relation was observed for colon cancer (Stable 1b). An increased risk was observed for lung cancer in most subgroups (Stable 4b). No association was observed for esophageal cancer, stomach cancer, pancreatic cancer, renal cancer, bladder cancer, breast cancer, ovarian cancer or melanoma in subgroup analyses (Stable 1b, 2b, 3b and 4b).

We conducted a sensitivity analysis by excluding studies that were not dose-response analyses to evaluate whether the results were stable. The summary relative risk for the highest versus lowest intake of coffee and liver cancer was 0.53 (95% CI = 0.43–0.65, I² = 0%, P = 0.693). The results were 0.87 (95% CI = 0.82–0.92, I² = 31.8%, P = 0.154), 0.70 (95% CI = 0.63–0.78, I² = 0%, P = 0.663), 0.85 (95% CI = 0.71–1.03, I² = 0%, P = 0.396), 0.62 (95% CI = 0.41–0.94, I² = 64.3%, P = 0.024) and 0.89 (95% CI = 0.72–1.11, I² = 0%, P = 0.882) for prostate cancer, endometrial cancer, colon cancer, oral, pharynx cancer and melanoma. The results were similar to those of the primary analysis except for colon cancer and melanoma.

Discussion

Our meta-analysis supports an inverse association between coffee intake and oral, pharynx cancer, liver cancer, colon cancer, prostate cancer, endometrial cancer and melanoma and increased association for lung cancer. Besides, the linear inverse associations were observed for liver cancer, prostate cancer and endometrial cancer. However, no significant association was found with esophageal cancer, stomach cancer, rectal cancer, pancreatic cancer, renal cancer, bladder cancer, prostate cancer, breast cancer, ovarian cancer, lung cancer, melanoma, and lymphoma.

To date, our study is the most comprehensive to conduct meta-analyses on coffee intake and cancer. We can clearly determine the association between coffee and most types of cancers from our research. Our findings seem to be more stable because our analyses were based on prospective studies. We conducted subgroup analyses and sensitivity analyses to explore the source of heterogeneity and verify the stability of the results. We also quantified the association between coffee intake and liver cancer, prostate cancer and endometrial cancer by conducting linear and non-linear dose-response analyses. Begg’s and Egger’s methods were used to evaluate publication bias. When the results suggested publication bias, we conducted a trim and fill analysis to determine whether the results were different.

Some limitations of our study should be taken into consideration. First, not all of the included studies conducted analyses by potential confounding, which could affect the results. However, we conducted subgroup analyses by confounding factors and found no substantial factors influencing the results. Second, some eligible studies did not provide relative risks and confidence intervals; thus, we extracted raw data for further analysis, which may influence the accuracy of the overall results regardless of confounding factors. Third, some studies measured the level of coffee intake in milliliters and others reported as cups. Due to the ambiguity in the volume contained in a cup of coffee, it was difficult to precisely compare the amount of coffee intake, which could influence the results of dose-response analyses. Moreover, we excluded a few studies for dose-response analyses. However, the results are similar to those of the analyses of highest versus lowest intake when we conducted sensitivity analyses by excluding these same data. Fourth, some cancers such as oral, pharynx cancer, esophageal cancer, renal cancer, lung cancer, melanoma and lymphoma included relatively limited number of studies for meta-analysis, may cause problems for evaluation of heterogeneities and publication bias and finally reduce the confidence of the results. Finally, our study did not include all types of cancers such as skin cancer, laryngeal carcinoma and glioma.

In addition, we found high heterogeneities when we conducted meta-analyses on coffee intake and oral, pharynx cancer and lung cancer. Although an inverse association was found between coffee and oral and
pharynx cancer, we did not conduct a further dose-response analysis because of substantial heterogeneity. The
dose-response analyses were also not conducted for colon cancer, melanoma and lung cancer because of the null
results of sensitivity analyses. Medium heterogeneities were found in the analyses of coffee intake and renal can-
cer, bladder cancer and stomach cancer. Heterogeneity contributes to the variance within included studies. The
biology diversity and methodological diversity could bring about heterogeneity. To further explore the source of
heterogeneity, we conducted subgroups analyses and meta-regression analyses based on many factors, such as
cancer subtypes, geographical location, sex, duration of follow-up and potential confounding factors. We found
publication bias in meta-analyses on coffee intake and liver cancer, bladder cancer and ovarian cancer. We con-
ducted trim and fill analyses and found that all of the results were stable. However, the evaluation for publication
bias is not very reliable when the number of include studies was small. Therefore, we could not draw conclusion
that there was no obvious publication bias for esophageal cancer, renal cancer, melanoma and lymphoma.

Coffee intake could reduce the incidence of colon cancer, liver cancer, prostate cancer, endometrial cancer,
oral, pharyngeal cancer and melanoma. There are several mechanisms attempting to explain this phenomenon.
Coffee contains many bioactive components, including caffeine, cafestol, kahweol, and chlorogenic acid. Some
studies indicate that caffeine can prevent oxidative DNA damage, modify the apoptotic response and reverse
the cell cycle checkpoint function. Moreover, some researchers have found that cafestol and kahweol are anticarcinogenic. Feng et al. hypothesized that chlorogenic acids can clear away reactive oxygen species and confer an anti-tumor effect. Dong et al. demonstrated that caffeine suppresses the progression of HCC through the Akt signaling pathway. Furthermore, caffeine and other compounds in coffee increased the clearance of estradiol and inhibited estradiol-mediated carcinogenesis in endometrial cells. Coffee consumption decreased the exposure of epithelial cells to carcinogens in the colon by increasing colonic motility. In addition, coffee has been reported to reduce the synthesis and secretion of bile acids, potential promoters of colon carcinogenesis.

The decreased risk of these types of cancers may attribute to some potential mechanism, however, there are many
controversies about the mechanisms. Although a positive association was observed between coffee intake and
cancer subtypes, geographical location, sex, duration of follow-up and potential confounding factors.

We found

References


57. Bidel, S.
55. Lee, K. J.
54. Oba, S.


45. Zamora-Ros, R.


145. Leviton, A. & Cowan, L. A review of the literature relating caffeine consumption by women to their risk of reproductive hazards.

Acknowledgements
We thank for their help Nature Publishing Group Language Editing in English language revision of this manuscript. We also thank for the support of International Science and Technology Cooperation Projects (2015DFA30650 and 2010DFB33720), Capital Special Research Project for Health Development (2014-2-4012), Capital research project for the characteristics clinical application (Z151100004015170) and Program for New Century Excellent Talents in University (NCET-11-0288).

Author Contributions
A.W. designed the study and wrote this manuscript. S.W. and C.Z. searched database and reviewed studies. C.H., C.W., X.W., X.Y., H.Z., Y.M. and L.H. collected and analyzed data. X.S. and H.Z. coordinated and provided financial support for this work.

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.


This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© The Author(s) 2016