Coffee consumption and the risk of cancer in the Norwegian Women and Cancer (NOWAC) study

Marko Lukic[1*] Idlir Licaj[1] Eiliv Lund[1] Guri Skeie[1] Elisabete Weiderpass[1234] Tonje Braaten[1]

[1]Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, The Arctic University of Norway, Tromsø, Norway

[2] Department of Research, Cancer Registry of Norway, Institute of Population-Based Cancer Research, Oslo, Norway.

[3] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[4] Genetic Epidemiology Group, Folkhälsan Research Center, Helsinki, Finland

Corresponding author*

marko.lukic@uit.no

Institutt for Samfunnsmedisin,

UiT Norges Arktiske Universitet, N - 9037 Tromsø, Norway

Tel: +47 776 20823; Mobil: +47 40 591581

Abstract

Background An association between coffee consumption and cancer has long been investigated. Coffee consumption among Norwegian women is high, thus this is a favorable population in which to study the impact of coffee on cancer incidence.

Methods Information on coffee consumption was collected from 91 767 women at baseline in the Norwegian Women and Cancer Study. These information were applied until follow-up information on coffee consumption, collected 6-8 years after baseline, became available. Multiple imputation was performed as a method for dealing with missing data in the cohort. Multivariable Cox regression models were used to calculate hazard ratios (HR) for breast, colorectal, lung, and ovarian cancer, as well as cancer at any site.

Results We observed a 17% reduced risk of colorectal cancer (95%CI 0.70-0.98, p_{trend}=0.10) and a 9% reduced risk of cancer at any site (95%CI 0.86-0.97, p_{trend}=0.03) in women who drank more than 3 and up to 7 cups/day, compared to women who drank ≤1 cups/day. A significantly increased risk of lung cancer was observed with a coffee consumption of >7 cups/day (HR=2.01, 95%CI 1.47-2.75, p_{trend}<0.001). This was most likely caused by residual confounding due to smoking, as no statistically significant association was observed in never smokers (>5 cups/day HR=1.42, 95%CI 0.44-4.57, p_{trend}=0.30). No significant association was found between coffee consumption and the risk of breast or ovarian cancer.

Conclusions In this study, coffee consumption was associated with a modest reduced risk of cancer at any site. Residual confounding due to smoking may have contributed to the positive association between high coffee consumption and the risk of lung cancer.

Key words: cancer, coffee, breast, colorectal, lung, ovarian, women, multiple imputation, prospective cohort study

Introduction

The Nordic countries lead the world in coffee consumption. Norway ranks second among them, with an average consumption of 9.4 kg/year per capita between 1997 and 2011, just behind Finland (11.7 kg), and ahead of Denmark (8.9 kg) and Sweden (8.1 kg) (1). Therefore, any causal association between coffee consumption and chronic diseases would have a significant public health impact in these countries.

We aimed to investigate the relationship between coffee consumption and the risk of breast, colorectal, ovarian, and lung cancers, as well as cancer at any site, in the Norwegian Women and Cancer (NOWAC) Study using baseline and follow-up information on total coffee consumption.

Results from the most recent meta-analysis suggest that high coffee consumption might be associated with a lower risk of colorectal cancer, and breast cancer in postmenopausal women (2;3). Moreover, a 27% increased risk of lung cancer was found for the highest coffee consumption group in a meta-analysis that combined the results of 13 studies, with a borderline non-significant inverse association being observed among never smokers (4). No significant relationship has been reported between coffee consumption and ovarian cancer (5). Overall, it seems that coffee might have a protective effect against cancer, as reported in a meta-analysis of 40 prospective cohort studies by Yu et al (6). In Norway, Stensvold and Jacobsen found a non-significant inverse association between coffee consumption and colon and rectal cancers in women, and a non-significant, increased risk of breast and lung cancers (7).

Breast, colorectal, and lung cancer are three of the most frequently diagnosed cancers in both Norway and worldwide (8;9). Ovarian cancer was included in order to complement the

71 study by Gavrilyuk et al (10) on coffee consumption and the risk of gynecologic tumors in the
72 NOWAC Study.

73 ## Methods

74 **The NOWAC study**

75 Detailed information on the NOWAC Study is available elsewhere (11). In short, random
76 samples of Norwegian women aged 30-70 years were invited to participate. More than 172
77 000 accepted and completed a questionnaire regarding their lifestyle, diet, and health status
78 (overall response rate: 52.7%). All women gave written informed consent. The NOWAC Study
79 was approved by the Regional Committee for Medical Research Ethics and the Norwegian
80 Data Inspectorate.

81 The cohort follow-up was conducted between 1996 and 2013. The baseline information
82 in this analysis were taken from the questionnaires of women enrolled in 1991-1992, 1996-
83 1997, 2003, and 2004. These women completed baseline food frequency questionnaires in
84 1998, 1996-1997, 2003, and 2004, respectively. We chose not to use the information
85 collected during the first wave of data collection (1991-1992) as the version of
86 questionnaires that was sent out did not include questions regarding diet. We decided to use
87 the information from the questionnaires sent in 1998 (the second wave of data collection)
88 for those women enrolled in the NOWAC from 1991-1992 as baseline data for the present
89 study. The information on coffee consumption was available for 98 405 women.

90 We excluded women with prevalent cancer other than non-melanoma skin cancer at
91 baseline and those who emigrated or died before the start of follow-up (N=4395), those who
92 were diagnosed with cancer after they emigrated (N=9), and those with total energy intake
93 above 15 000 kJ or below 2500 kJ per day (N=619). Finally, we excluded 1615 women that
94 had missing information on coffee consumption at baseline, i.e. the women who did not

answer to none of the three questions regarding boiled, instant and filtered coffee intake in the first questionnaire. Thus, the final analytical study sample consisted of 91 767 women. Follow-up information were collected from 79 461 of these women, who received the follow-up questionnaire before the end of the study, 6-8 years after baseline data collection. The rest of the women (N=12 306) received the baseline questionnaire in 2004, while the follow-up questionnaire was sent out to them after the present study has ended.

Assessment of coffee consumption and covariates

Women answered the same question on coffee consumption at baseline and at follow-up: "How many cups of each kind of coffee (boiled, filtered, instant) did you usually drink during the past year?" Women could choose from the following answers: never/seldom, 1-6 cups/week, 1 cup/day, 2-3 cups/day, 4-5 cups/day, 6-7 cups/day, and ≥8 cups/day for each brewing method. Total coffee consumption was derived by summing the frequencies of each of the brewing methods and was categorized as ≤1 cup/day (light consumers), more than 1 up to 3 cups/day (low moderate consumers), more than 3 up to 7 cups/day (high moderate consumers), and >7 cups/day (heavy consumers). As the size of a cup was not specified in the questionnaire, 2.1 dl was used as the standard cup size (12).

Women also answered questions on smoking status (never, former, or current), and number of pack-years (calculated as number of cigarettes smoked/day divided by 20 and multiplied by years of smoking) at baseline and at follow-up. Women who reported they were current or former smokers at baseline and never smokers at follow-up were categorized as former smokers at follow-up (N=1608). Additionally, the information on BMI, physical activity, alcohol consumption, total energy intake, and use of hormone replacement therapy (never, former, current) were also collected both at baseline and follow-up.

Cancer incidence, death, and emigration

Information on cancer incidence, death, and emigration in the cohort was obtained through linkage to the Norwegian Cancer Registry, the Cause of Death Registry, and the Norwegian Central Population Register, respectively, using the unique 11-digit personal number assigned to every legal resident in Norway. The 7th Revision of the International Statistical Classification of Diseases, Injuries and Causes of Death was used to classify breast (170.0-170.9), colorectal (153.0-154.0), ovarian (175.0-175.9), and lung (162.0-162.1) cancer cases in the Cancer Registry of Norway.

Statistical methods

As per the methods proposed by Hu et al (13), we applied baseline information until follow-up information became available, until date of diagnosis of any incident cancer other than non-melanoma skin cancer, death, or emigration, whichever occurred first. Thereafter follow-up information was applied until diagnosis of any incident cancer other than non-melanoma skin cancer, until death, emigration or the end of the study period (31 December 2013), whichever occurred first.

Cox proportional hazards regression models were used to calculate hazard ratios (HRs) for developing breast, colorectal, ovarian, or lung cancer, as well as cancer at any site other than non-melanoma skin cancer, with 95% confidence intervals (CIs) for each coffee consumption group. Light consumers (i.e., those drinking ≤1 cup/day), were used as the reference group, as it was impossible to differentiate between coffee abstainers and occasional coffee drinkers from the answers offered in the questionnaire. Attained age was used as the underlying time scale. All models were stratified by questionnaire subcohorts in order to control for potential differences in the long follow-up time.

142 We decided to use follow-up information on smoking exposure in addition to coffee
143 consumption, for both complete-case analyses and analyses performed on multiple imputed
144 datasets. This was done as the prevalence of current smokers varied over time in the cohort,
145 and as we suspected a strong confounding effect of the smoking exposure in the analyses.
146 Analyses for each cancer site were adjusted for known risk factors (9) in the preliminary,
147 complete-case analysis, which included baseline information only. The preliminary models
148 for each cancer site were adjusted for a selection of the following covariates: menopausal
149 status (premenopausal/postmenopausal), smoking status (never, former, current), age at
150 smoking initiation (<20, ≥20 years), number of pack-years (≤14, 15-19, ≥20), exposure to
151 cigarette smoke during childhood (yes/no), duration of education (≤9, 10-12, 13-16, ≥17
152 years), body mass index (BMI, ≤18.49, 18.5-24.9, 25-29.9, and ≥30 kg/m^2), physical activity
153 level (1-4, 5-6, 7-10), alcohol consumption (0, 0.1-3.99, 4-9.99, ≥10 g/day), number of
154 children (0, 1-2, ≥3), age at first birth (<20, 20-24, 25-29, ≥30 years), ever use of oral
155 contraceptives (yes/no), duration of oral contraceptive use in years (continuous), use of
156 hormone replacement therapy (never, former, current), maternal history of breast cancer
157 (yes/no), total energy intake (tertiles, kJ/day), intake of fibers (<=20, >20 g/day), intake of
158 processed meat (continuous, g/day), intake of red meat (<=10, 10.01-20, >20, g/day), height
159 (continuous, cm), and participation in mammography screening (yes/no). In order to be
160 retained in the final model, the removal of the covariate had to lead to a change in the
161 regression coefficients of at least 10% in any of the coffee consumption groups.
162 If a linear trend was observed for a specific covariate, that covariate was treated as
163 continuous. When the adjustment required all the smoking variables in the analysis, we
164 modelled these as five categorical variables, which included the information on smoking
165 status, age at smoking initiation, and number of pack-years. Similarly, 12 categorical

166 variables were made by combining the information on number of children and age at first
167 birth for the breast cancer analysis. As in other large cohort studies, when age at menopause
168 was not available, the age 53 years was used as the threshold by which to classify
169 premenopausal and postmenopausal women in the complete-case analyses (14).
170 An interaction between coffee consumption and the logarithmic transformation of
171 participants' age was tested to check the proportional hazards assumption. To test for linear
172 trend, a median value was assigned to each category of ordinal coffee consumption variable,
173 which was then modeled as continuous in the analyses. We assessed possible interactions
174 between coffee consumption and smoking status, BMI, and physical activity level,
175 respectively, as these had the potential to interact with the antioxidant effects of coffee, or
176 could affect the metabolism of coffee compounds (15-17).
177 In order to counteract residual confounding due to smoking, we repeated the analysis on
178 lung cancer using women that were never smokers during the entire study period. For this
179 analysis, we classified heavy consumers as those drinking >5 cups/day in order to increase
180 number of cases in the highest coffee consumption category. We have also conducted
181 complete-case sensitivity analyses in which we have used, depending on the outcome, the
182 follow-up information on BMI, physical activity, alcohol consumption, total energy intake, or
183 use of hormone replacement therapy. As a complementary analyses, we conducted an
184 analysis for colon and rectal cancers separately. We repeated the analyses for each of the
185 outcomes after excluding cancers at the corresponding sites diagnosed during the first two
186 years of follow-up in order to control for possible reverse causality. Furthermore, we did the
187 analyses in which we had excluded cancer cases of interest that occurred during the first
188 year of follow-up, and at the same time censoring at the time of answering the second

questionnaire those cancer cases diagnosed during the first year after they received the second questionnaire.

Multiple imputation

Under the assumption that data was missing at random, and after confirming that the pattern of missingness was arbitrary, we performed multiple imputation to deal with missing information at baseline and follow-up. Twenty duplicate datasets were created in order to reduce sampling variability from the imputation simulation (18). The missing values from baseline and follow-up were then replaced by imputed values based on the observed information. Separate imputation models were created for each outcome, including all of the variables from the final analysis of the specific cancer sites. In addition, in order to increase the predictive power of the imputation procedure, we included smoking status and number of pack-years (baseline and follow-up information), and age at smoking initiation, duration of education, BMI, physical activity level, and alcohol consumption (baseline information) in each imputation model, regardless of whether the variable(s) were used in the multivariable Cox regression model.

In order to avoid possible inconsistencies, we imputed the "change in smoking status" between baseline and follow-up. Later, we used these imputed values to determine if a person was a never, former, or current smoker at follow-up. Similarly, we imputed the difference in the number of pack-years between baseline and follow-up, in order to avoid lower imputed values at follow-up compared to baseline.

If the interaction term between coffee consumption and any one of the variables smoking status, BMI, or physical activity level was statistically significant in the complete-case analysis, these terms were included as predictors in the imputation model. We also

used the Nelson-Aalen cumulative hazard estimator as a predictor in all the imputation models (19).

The estimates from the twenty imputed datasets were combined using Rubin's rules in order to obtain HRs and corresponding 95% CIs (20). All the analyses and the multiple imputations were done in STATA version 14.0 (Stata Corp, College Station, TX, USA).

Results

During an average of 13.1 years of follow-up and 1.2 million person-years, 9675 cases of cancer were diagnosed: 3277 (33.9%) breast cancers, 1266 (13.1%) colorectal cancers, 446 (4.6%) ovarian cancers, and 819 (8.5%) lung cancers. The ten most common cancer sites in the NOWAC study are presented in Supplementary table 1. At baseline, most women reported they were high moderate consumers (more than 3 up to 7 cups/day; 42.8%). At follow-up, the proportion of high moderate consumers and heavy consumers (>7 cups/day) decreased. Distribution of participants according to filtered, instant, and boiled coffee consumption at baseline and follow-up is presented in Supplementary table 2.

The proportion of women diagnosed with cancer at any site was largest among heavy consumers (12.4%); this was also the case for lung cancer (2.4%). Light consumers (≤1 cup/day) were more likely to have fewer children, were the oldest at the time of first birth, were more likely to have used oral contraceptives, and had the lowest energy intake compared to women in other coffee consumption groups. Heavy consumers were the youngest at baseline, had the highest BMI, and the lowest physical activity level score. In addition, these women consumed less alcohol, had more children, were younger at first birth, and were less likely to have used hormone replacement therapy compared to women in other coffee consumption groups (Table 1).

The proportion of current smokers was the lowest among light coffee consumers, and became higher in each subsequent coffee consumption category, with the percentage among heavy consumers reaching 68.5%. A positive relationship was also observed between both number of pack-years and age at smoking initiation, and the number of cups/day of coffee consumed. In contrast, we found a negative trend for coffee consumption and duration of education, with light consumers averaging 13 years of school, compared to the 10.6 years observed among heavy consumers (Table 1).

The highest proportion of missing values was observed for age at menopause at baseline and follow-up (54.4%), coffee consumption at follow-up (27.0 %), smoking status at follow-up (27.4 %), and number of pack-years at follow-up (42.8 %). The highest proportion of missing information on coffee consumption at follow-up was observed among those who reported being light consumers at baseline. Women that were heavy coffee consumers were more likely to have missing information on smoking at both baseline and follow-up (Table 2). The comparison between the complete-case dataset and the dataset with imputed values are presented in Supplementary Table 3. The results of the complete-case analysis for each of the outcomes are reported in Supplementary Table 4.

The following results are those from the analyses performed on the imputed datasets. We observed a 9% reduction in the risk of cancer at any site among high moderate consumers compared to light coffee consumers (HR=0.91, 95% CI 0.86-0.97, p_{trend}=0.03) (Table 3). There was no significant association between coffee consumption and the risk of breast cancer when heavy consumers were compared with the reference group (HR=0.87, 95% CI 0.71-1.06, p_{trend}=0.06). A borderline non-significant HR of 0.91 was found among high moderate consumers (95% CI 0.82-1.00). A statistically significant inverse association between coffee consumption and the risk of colorectal cancer was found only in high

moderate consumers, with no significant linear trend (HR=0.83, 95% CI 0.70-0.98, p_{trend}=0.10). No association was found between coffee consumption and the risk of ovarian cancer (highest vs. lowest consumption category HR=0.87, 95%CI 0.50-1.51, p_{trend}=0.89). Compared to light consumers, heavy consumers had a more than five-fold higher risk of lung cancer in the age-adjusted analysis. (95% CI 4.20-7.60). This association was attenuated after multivariable adjustment, but an increase in risk was still observed in the highest coffee consumption group (HR=2.01, 95% CI 1.47-2.75, p_{trend}<0.001) (Table 3).

We found no statistically significant association between coffee consumption and the risk of lung cancer among never smokers (HR=1.42 among women who drank >5 cups/day, 95% CI 0.44-4.57, p_{trend}=0.30) (Table 4).

None of the interactions tested between coffee consumption and smoking status, BMI, and physical activity level were significant in any of the outcomes investigated (data not shown). We found no interaction effect between coffee consumption and the logarithmic transformation of age in any of the outcomes.

The risk estimates for, colorectal, ovarian, lung and cancer at any site from the lag analyses were similar to those from the analyses that included the entire study sample (results not shown). However, we observed a significantly decreased risk of breast cancer for low and high moderate coffee consumers after we excluded breast cancer cases diagnosed during the first two years of follow-up (HR=0.90, 95% CI 0.81-0.99; HR=0.86, 95% CI 0.78-0.96, p_{trend}=0.01).

The complete-case analyses in which follow-up information on BMI, physical activity, alcohol consumption, total energy intake and use of hormonal replacement therapy were used in addition to coffee and smoking exposure variables, revealed similar results with the analyses in which only coffee and smoking variables were updated (results not shown).

Finally, in both complete-case and the analyses on multiple imputed datasets performed for colon and rectal cancers separately, we found no evidence of an association between coffee consumption and either colon or rectal cancer risk (Supplementary table 5).

The associations between coffee consumption and the risk of breast, colorectal, lung, and ovarian cancer, as well as cancer at any site among never smokers are presented in Supplementary table 6.

Discussion

We observed a decreased risk of colorectal cancer and of cancer at any site associated with high moderate coffee consumption, with no evidence of linear relationship between coffee consumption and colorectal cancer risk. In contrast, we found a statistically significant association between high coffee consumption (>7 cups/day) and the risk of lung cancer. However, no significant association between coffee intake and the risk of lung cancer was observed in never smokers.

The main strengths of our study include its prospective design, the relatively large sample size, and the statistical power necessary to detect differences between the coffee consumption groups in each of the studied cancer sites. The participants in the NOWAC cohort were randomly recruited from the general population. The external validity of NOWAC study has been previously found to be acceptable. Briefly, the response rate from the NOWAC study is similar to many other populated-based cohorts. The authors found that the responders do not differ materially from the source population except for somewhat higher educational level. Similarly, the observed incidence rates for all cancer sites in the NOWAC study were comparable to national figures (21). Linkage to the Norwegian Cancer Registry via the unique person number allowed us to obtain virtually complete follow-up.

The food frequency questionnaires used in the NOWAC Study were validated by 24-h dietary recalls study (12), which showed a high validity of information on coffee consumption (Spearman's correlation coefficient r=0.82). We used repeated measurements of coffee consumption and smoking exposure in order to take into account changes in these variables over time and to attenuate the risk of measurement error. Moreover, the use of the updated information on coffee consumption allowed us to conduct an extensive lag analysis in order to check for possible reverse causality. Finally, we used multiple imputation to maximize the number of participants and cancer cases included the analyses.

There are also several limitations in our study. We lacked power to explore the risk of some cancer sites such as liver that were found to be inversely associated with coffee intake. The risk of hepatocellular carcinoma was previously reported to be lower in the higher categories of coffee consumption (22;23). During the follow-up, 44 women were diagnosed with primary liver and biliary passages cancer in the present study. Any analysis with this low number of cases would lead to unreliable results. We did not have information regarding caffeination status. However, the consumption of decaffeinated coffee is very uncommon in Norway. We did not conduct a separate analysis for different brewing types of coffee, as the number of women that reported drinking more than 7 cups of instant or boiled coffee at baseline was low (213 and 999, respectively). As the consumption of boiled coffee is decreasing in the cohort, the number of participants in the highest coffee consumption category was not sufficient for analyses of either of these brewing types. We believe, however, that our results were driven by filtered coffee, which was the most commonly consumed among women in the cohort.

The effect of residual confounding cannot be excluded, although we adjusted for many known risk factors. This may particularly be the case for the association between heavy

coffee consumption and the risk of lung cancer, which is most likely due to residual confounding from smoking. Indeed, there were pronounced differences in coffee consumption between never, former, and current smokers in the cohort. As the proportion of daily smokers in Norway is decreasing (24), adjusting for only baseline information on smoking exposure could have yielded biased estimates. However, we lacked the information necessary to adjust for more comprehensive markers of smoking exposure, such as if a person inhales smoke from a cigarette, or lifetime exposure to secondhand smoke and other pollutants. Tea consumption was not taken into account in the analyses, as this information was not available from the NOWAC questionnaires. Therefore, a possible confounding effect of tea, which contains some of the same bioactive components as coffee, cannot be excluded.

Although the information on coffee consumption was shown to be valid based on the results from the validation study, misclassification is still possible. We tried to reduce within-person variation and minimize the risk of misclassification bias by using follow-up information. However, as coffee consumption was self-reported, misclassification cannot be completely ruled out.

We decided to impute missing information at baseline and follow-up, assuming a missing-at-random mechanism. We introduced a wide range of variables into the imputation models, which we thought could be used to predict incomplete variables or to predict whether the incomplete variable was missing (25). However, it is possible that at least some of the information is still missing-not-at-random and thus that our estimates are not free of bias.

To our knowledge, this is the first study examining the effect of coffee consumption on the risk of cancer that used repeated information on coffee consumption and combined this method with multiple imputation of missing data.

The results from our study are in line with the meta-analysis by Yu et al regarding coffee intake and the overall risk of cancer, in which a 13% risk reduction was found in women. However, the study authors did not specify which coffee consumption group was compared to the non/lowest drinking category (6). In a prospective study from Norway, which included 21 238 women, a non-significant inverse association was observed in the highest coffee consumption group (≥7 cups/day) (7). The results from the Swedish Västerbotten Intervention Project (VIP) cohort showed a non-significant HR of 0.92 for all cancer sites in both men and women who drank coffee on at least 4 occasions per day compared to the reference group (26).

Our findings regarding the risk of breast cancer are in accordance with the recent meta-analysis, as well as with the studies from Norway, France, Netherlands, and Sweden (3;7;27-29). In addition, the results from Nurses' Health Study, which included follow-up information on coffee consumption support our findings (≥4 cups/day HR=0.92 95% CI 0.82-1.03) (30). No significant associations were also found between total coffee consumption and the risk of breast cancer in pre- or postmenopausal women in the EPIC study (31). On the other hand, another study from Sweden found a significant 19% decrease in risk among women who drank at least 5 cups of coffee per day (32). We did, however, find a similar risk reduction for the women drinking more than 3 and up to 5 cups of coffee/day, after we excluded breast cancer cases that were diagnosed during the first two years after enrollment in the study.

Our results regarding the risk of colorectal cancer depart somewhat from the findings of other cohort studies. Studies from Sweden and the United States also utilized updated information on coffee consumption, but found no association between high coffee consumption and the risk of colorectal cancer in women (33;34). Authors from the EPIC cohort also concluded that coffee consumption was not likely to be associated with the risk of colorectal cancer, as did the authors of the Japan Collaborative Cohort Study for the Evaluation of Cancer Risk (35;36). In the most recent meta-analysis, a significant inverse association was found in women after pooling the results from 25 case-control studies (summary OR=0.82). However, no such findings were found in the meta-analysis that included cohort studies (2). However, even though we found an association between high moderate coffee consumption and colorectal cancer risk, an absence of a linear relationship supports the findings from the mentioned studies. Furthermore, coffee consumption was associated with neither colon nor rectal cancer in the separate analyses.

The observed differences in the results regarding the risk of colorectal cancer might be due to differences in the potential confounders that were taken into account in the analyses. Indeed, the only study that carried out a detailed adjustment for smoking exposure that was comparable to ours was the EPIC study. Lack of adjustment for family history of colorectal cancer, the information not available for our cohort, could partially explain the differences between our study results and those from Japan, Sweden, and the United States.

Our findings regarding ovarian cancer are in agreement to those from the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, and from a meta-analysis by the same authors, both of which showed no association between high coffee consumption and the risk of ovarian cancer (5).

In an updated meta-analysis of epidemiological studies, Xie et al found a significant positive association between high coffee consumption and lung cancer in men, while a non-significant summary OR of 1.16 was observed for women in the highest coffee consumption category (37). One possible explanation for the difference between the meta-analysis and the present study was that Xie et al used 3 cups/day as the cut-off between the moderate and highest coffee consumption group. Similarly, in a recent study from the United States, a non-significant higher risk of lung cancer in women was associated with the highest level of coffee intake, defined as ≥4 cups/day (HR=1.10; 95% CI 0.95-1.26) (38). Finally, a study from Norway found a two-fold increased risk of lung cancer in women that were consuming at least 7 cups of coffee per day (7). Residual confounding by smoking is likely to have influenced the effect estimates in our study, as well as in previous studies. A strong correlation between smoking habits and coffee consumption can be at least partially explained by the fact that caffeine and nicotine share a metabolic pathway, via the CYP1A2 gene (39;40). It seems that an analysis with a sub-optimal adjustment for smoking exposure would likely yield a positive association between coffee consumption and the risk of lung cancer. This is also supported by the lack of statistically significant association we observed between coffee consumption and risk of lung cancer among never smokers, as was also found in the study by Guertin et al (38). Our results among never smokers are in line with two meta-analyses in which no significant associations were observed between coffee intake and the risk of lung cancer (4;37). However, an inverse association reported in the meta-analyses contradict the positive association in the present study. Our analysis was, however, hampered by a small number of lung cancer cases among never smokers. As the result, the interpretation of these results warrants some caution.

A number of biologically active substances contained in roasted coffee have the potential to either suppress or induce carcinogenesis. Chlorogenic acid is one of the ingredients that contributes significantly to the antioxidant effect of coffee. It has been hypothesized that chlorogenic acid could alter the risk of some cancers by reducing glucose levels in the blood and increasing insulin sensitivity (41;42). Kahweol, one of the diterpenes that constitutes coffee, has been found to induce apoptosis in human leukemia cells (43), to reduce gentoxicity in hepatoma cells (44), and to induce synthesis of endogenous antioxidants (45). Caffeine has also been shown to alter the risk of malignancies in pre- and postmenopausal women by increasing the level of sex-hormone binding globulin and decreasing the levels of free estradiol (46).

Even though the observed positive association between coffee consumption and the risk of lung cancer is likely due to residual confounding from smoking, we cannot rule out the possibility of a biological effect of some coffee compounds on lung cancer. The adverse effects of caffeine are mainly related to its ability to inhibit DNA repair mechanisms (47;48). Muller et al argued that caffeine negatively effects both the speed of DNA repair, and the residual damage after exposing mammalian cells to radiation (49).

Conclusion

The results from our study indicate that high moderate coffee intake may have a protective effect on the overall risk of cancer. The observed positive association between heavy coffee consumption and the risk of lung cancer should be interpreted with caution, as residual confounding due to smoking exposure is probable.

Authors' contributions:

ML carried out the statistical analysis and drafted the manuscript. IL contributed with the interpretation of the data and revision of the manuscript. EL was responsible for critical revision of the manuscript. EL is also the PI of the NOWAC. GS and EW critically revised the manuscript. TB developed the research plan, prepared the data, revised the manuscript, and provided critical help for the multiple imputation modeling.

Reference List

(1) International Coffee Council. Trends in coffee consumption in selected importing countries. 2015.

(2) Li G, Ma D, Zhang Y, Zheng W, Wang P. Coffee consumption and risk of colorectal cancer: a meta-analysis of observational studies. Public Health Nutr 2013 Feb;16(2):346-57.

(3) Jiang W, Wu Y, Jiang X. Coffee and caffeine intake and breast cancer risk: an updated dose-response meta-analysis of 37 published studies. Gynecol Oncol 2013 Jun;129(3):620-9.

(4) Tang N, Wu Y, Ma J, Wang B, Yu R. Coffee consumption and risk of lung cancer: a meta-analysis. Lung Cancer 2010 Jan;67(1):17-22.

(5) Braem MG, Onland-Moret NC, Schouten LJ, Tjonneland A, Hansen L, Dahm CC, et al. Coffee and tea consumption and the risk of ovarian cancer: a prospective cohort study and updated meta-analysis. Am J Clin Nutr 2012 May;95(5):1172-81.

(6) Yu X, Bao Z, Zou J, Dong J. Coffee consumption and risk of cancers: a meta-analysis of cohort studies. BMC Cancer 2011;11:96.

(7) Stensvold I, Jacobsen BK. Coffee and cancer: a prospective study of 43,000 Norwegian men and women. Cancer Causes Control 1994 Sep;5(5):401-8.

(8) Cancer Registry of Norway IoPCR. Cancer in Norway 2013. 2015.

(9) World Cancer Research Fund, American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. 2007.

(10) Gavrilyuk O, Braaten T, Skeie G, Weiderpass E, Dumeaux V, Lund E. High coffee consumption and different brewing methods in relation to postmenopausal endometrial cancer risk in the Norwegian women and cancer study: a population-based prospective study. BMC Womens Health 2014;14:48.

(11) Lund E, Dumeaux V, Braaten T, Hjartaker A, Engeset D, Skeie G, et al. Cohort profile: The Norwegian Women and Cancer Study--NOWAC--Kvinner og kreft. Int J Epidemiol 2008 Feb;37(1):36-41.

(12) Hjartaker A, Andersen LF, Lund E. Comparison of diet measures from a food-frequency questionnaire with measures from repeated 24-hour dietary recalls. The Norwegian Women and Cancer Study. Public Health Nutr 2007 Oct;10(10):1094-103.

(13) Hu FB, Stampfer MJ, Rimm E, Ascherio A, Rosner BA, Spiegelman D, et al. Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. Am J Epidemiol 1999 Mar 15;149(6):531-40.

(14) Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet 2003 Aug 9;362(9382):419-27.

(15) Benowitz NL, Peng M, Jacob P, III. Effects of cigarette smoking and carbon monoxide on chlorzoxazone and caffeine metabolism. Clin Pharmacol Ther 2003 Nov;74(5):468-74.

(16) Bakuradze T, Boehm N, Janzowski C, Lang R, Hofmann T, Stockis JP, et al. Antioxidant-rich coffee reduces DNA damage, elevates glutathione status and contributes to weight control: results from an intervention study. Mol Nutr Food Res 2011 May;55(5):793-7.

(17) Collomp K, Anselme F, Audran M, Gay JP, Chanal JL, Prefaut C. Effects of moderate exercise on the pharmacokinetics of caffeine. Eur J Clin Pharmacol 1991;40(3):279-82.

(18) Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.

(19) White IR, Royston P. Imputing missing covariate values for the Cox model. Stat Med 2009 Jul 10;28(15):1982-98.

(20) Rubin DB. Multiple imputation after 18+ years. Journal of the American Statistical Association 1996 Jun;91(434):473-89.

(21) Lund E, Kumle M, Braaten T, Hjartaker A, Bakken K, Eggen E, et al. External validity in a population-based national prospective study--the Norwegian Women and Cancer Study (NOWAC). Cancer Causes Control 2003 Dec;14(10):1001-8.

(22) Bamia C, Lagiou P, Jenab M, Trichopoulou A, Fedirko V, Aleksandrova K, et al. Coffee, tea and decaffeinated coffee in relation to hepatocellular carcinoma in a European population: multicentre, prospective cohort study. Int J Cancer 2015 Apr 15;136(8):1899-908.

(23) Aleksandrova K, Bamia C, Drogan D, Lagiou P, Trichopoulou A, Jenab M, et al. The association of coffee intake with liver cancer risk is mediated by biomarkers of inflammation and hepatocellular injury: data from the European Prospective Investigation into Cancer and Nutrition. Am J Clin Nutr 2015 Dec;102(6):1498-508.

(24) Statistics Norway. Smoking habits, 2014.

(25) White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011 Feb 20;30(4):377-99.

(26) Nilsson LM, Johansson I, Lenner P, Lindahl B, Van GB. Consumption of filtered and boiled coffee and the risk of incident cancer: a prospective cohort study. Cancer Causes Control 2010 Oct;21(10):1533-44.

(27) Hirvonen T, Mennen LI, de BA, Castetbon K, Galan P, Bertrais S, et al. Consumption of antioxidant-rich beverages and risk for breast cancer in French women. Ann Epidemiol 2006 Jul;16(7):503-8.

(28) Bhoo PN, Peeters P, van GC, Beulens JW, van der Graaf Y, Bueno-de-Mesquita B, et al. Coffee and tea intake and risk of breast cancer. Breast Cancer Res Treat 2010 Jun;121(2):461-7.

(29) Michels KB, Holmberg L, Bergkvist L, Wolk A. Coffee, tea, and caffeine consumption and breast cancer incidence in a cohort of Swedish women. Ann Epidemiol 2002 Jan;12(1):21-6.

(30) Ganmaa D, Willett WC, Li TY, Feskanich D, van Dam RM, Lopez-Garcia E, et al. Coffee, tea, caffeine and risk of breast cancer: a 22-year follow-up. Int J Cancer 2008 May 1;122(9):2071-6.

(31) Bhoo-Pathy N, Peeters PH, Uiterwaal CS, Bueno-de-Mesquita HB, Bulgiba AM, Bech BH, et al. Coffee and tea consumption and risk of pre- and postmenopausal breast cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort study. Breast Cancer Res 2015;17:15.

(32) Oh JK, Sandin S, Strom P, Lof M, Adami HO, Weiderpass E. Prospective study of breast cancer in relation to coffee, tea and caffeine in Sweden. Int J Cancer 2015 Apr 17.

(33) Larsson SC, Bergkvist L, Giovannucci E, Wolk A. Coffee consumption and incidence of colorectal cancer in two prospective cohort studies of Swedish women and men. Am J Epidemiol 2006 Apr 1;163(7):638-44.

(34) Michels KB, Willett WC, Fuchs CS, Giovannucci E. Coffee, tea, and caffeine consumption and incidence of colon and rectal cancer. J Natl Cancer Inst 2005 Feb 16;97(4):282-92.

(35) Dik VK, Bueno-de-Mesquita HB, Van Oijen MG, Siersema PD, Uiterwaal CS, Van Gils CH, et al. Coffee and tea consumption, genotype-based CYP1A2 and NAT2 activity and colorectal cancer risk-results from the EPIC cohort study. Int J Cancer 2014 Jul 15;135(2):401-12.

(36) Yamada H, Kawado M, Aoyama N, Hashimoto S, Suzuki K, Wakai K, et al. Coffee consumption and risk of colorectal cancer: the Japan Collaborative Cohort Study. J Epidemiol 2014;24(5):370-8.

(37) Xie Y, Qin J, Nan G, Huang S, Wang Z, Su Y. Coffee consumption and the risk of lung cancer: an updated meta-analysis of epidemiological studies. Eur J Clin Nutr 2015 Jun 17.

(38) Guertin KA, Freedman ND, Loftfield E, Graubard BI, Caporaso NE, Sinha R. Coffee consumption and incidence of lung cancer in the NIH-AARP Diet and Health Study. Int J Epidemiol 2015 Jun 16.

(39) Gu L, Gonzalez FJ, Kalow W, Tang BK. Biotransformation of caffeine, paraxanthine, theobromine and theophylline by cDNA-expressed human CYP1A2 and CYP2E1. Pharmacogenetics 1992 Apr;2(2):73-7.

(40) Hukkanen J, Jacob P, III, Peng M, Dempsey D, Benowitz NL. Effect of nicotine on cytochrome P450 1A2 activity. Br J Clin Pharmacol 2011 Nov;72(5):836-8.

(41) Shearer J, Farah A, de PT, Bracy DP, Pencek RR, Graham TE, et al. Quinides of roasted coffee enhance insulin action in conscious rats. J Nutr 2003 Nov;133(11):3529-32.

(42) Renehan AG, Roberts DL, Dive C. Obesity and cancer: pathophysiological and biological mechanisms. Arch Physiol Biochem 2008 Feb;114(1):71-83.

(43) Oh JH, Lee JT, Yang ES, Chang JS, Lee DS, Kim SH, et al. The coffee diterpene kahweol induces apoptosis in human leukemia U937 cells through down-regulation of Akt phosphorylation and activation of JNK. Apoptosis 2009 Nov;14(11):1378-86.

(44) Majer BJ, Hofer E, Cavin C, Lhoste E, Uhl M, Glatt HR, et al. Coffee diterpenes prevent the genotoxic effects of 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) and N-nitrosodimethylamine in a human derived liver cell line (HepG2). Food Chem Toxicol 2005 Mar;43(3):433-41.

(45) Higgins LG, Cavin C, Itoh K, Yamamoto M, Hayes JD. Induction of cancer chemopreventive enzymes by coffee is mediated by transcription factor Nrf2. Evidence that the coffee-specific diterpenes cafestol and kahweol confer protection against acrolein. Toxicol Appl Pharmacol 2008 Feb 1;226(3):328-37.

(46) Kotsopoulos J, Eliassen AH, Missmer SA, Hankinson SE, Tworoger SS. Relationship between caffeine intake and plasma sex hormone concentrations in premenopausal and postmenopausal women. Cancer 2009 Jun 15;115(12):2765-74.

(47) Deplanque G, Vincent F, Mah-Becherel MC, Cazenave JP, Bergerat JP, Klein-Soyer C. Caffeine does not cause override of the G2/M block induced by UVc or gamma radiation in normal human skin fibroblasts. Br J Cancer 2000 Aug;83(3):346-53.

(48) Deplanque G, Ceraline J, Mah-Becherel MC, Cazenave JP, Bergerat JP, Klein-Soyer C. Caffeine and the G2/M block override: a concept resulting from a misleading cell kinetic delay, independent of functional p53. Int J Cancer 2001 Nov 1;94(3):363-9.

(49) Muller WU, Bauch T, Wojcik A, Bocker W, Streffer C. Comet assay studies indicate that caffeine-mediated increase in radiation risk of embryos is due to inhibition of DNA repair. Mutagenesis 1996 Jan;11(1):57-60.