Blood polyphenol concentrations and differentiated thyroid carcinoma in women from the European Prospective Investigation into Cancer and Nutrition (EPIC) study.

Raul Zamora-Ros[1,*], Leila Lujan-Barroso[1], David Achaintre[2], Silvia Franceschi[3], Cecilie Kyrø[4], Kim Overvad[5], Anne Tjønneland[4,6], Therese Truong[7,8], Lucie Lecuyer[7,8], Marie-Christine Boutron-Ruault[7,8], Verena Katzke[9], Theron S. Johnson[9], Matthias B. Schulze[10,11], Antonia Trichopoulou[12], Eleni Peppa[12], Carlo La Vechia[12,13], Giovanna Masala[14], Valeria Pala[15], Salvatore Panico[16], Rosario Tumino[17], Fulvio Ricceri[18,19], Guri Skeie[20], J. Ramón Quirós[21], Miguel Rodriguez-Barranco[22,23,24,25], Pilar Amiano[24,26], María-Dolores Chirlaque[24,27], Eva Ardanaz[24,28,29], Martin Almquist[30], Joakim Hennings[31], Roel Vermeulen[32,33], Nicholas J. Wareham[34], Tammy Y.N. Tong[35], Dagfinn Aune[36,37,38], Graham Byrnes[2], Elisabete Weiderpass[2], Augustin Scalbert[2], Sabina Rinaldi[2], Antonio Agudo[1]

AUTHOR AFFILIATIONS:

[1]Unit of Nutrition and Cancer, Epidemiology Research Program, Catalan Institute of Oncology, Bellvitge Biomedical Research Institute (IDIBELL), Hospitalet de Llobregat (Barcelona), Spain.

[2]International Agency for Research on Cancer (IARC-WHO), Lyon, France

[3]Centro di Riferimento Oncologico (CRO), Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Aviano, Italy

[4]Danish Cancer Society Research Center, Copenhagen, Denmark

[5]Department of Public Health, Aarhus University, Aarhus, Denmark

[6]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[7]Université Paris-Saclay, UVSQ, Inserm, CESP, 94807, Villejuif, France.

[8]Institut Gustave Roussy, 94805, Villejuif, France

[9]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

[10]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany

[11]Institute of Nutritional Sciences, University of Potsdam, Nuthetal, Germany

[12]Hellenic Health Foundation, Athens, Greece

[13]Department of Clinical Sciences and Community Health Università degli Studi di Milano, Milan, Italy

[14]Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network - ISPRO, Florence, Italy

[15]Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy.

[16]Dipartimento di Medicina Clinica e Chirurgia Federico II University, Naples, Italy.

[17]Cancer Registry and Histopathology Department, "Civic - M.P. Arezzo" Hospital, ASP Ragusa, Italy.

[18]Department of Clinical and Biological Sciences,University of Turin, Turin, Italy

[19]Unit of Epidemiology, Regional Health Service ASL TO3, Grugliasco, Turin, Italy

[20]Department of Community Medicine, UiT the Arctic University of Norway, Tromsø, Norway.

[21]Public Health Directorate, Asturias, Spain.

[22]Escuela Andaluza de Salud Pública (EASP), Granada, Spain.

[23]Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain

[24]CIBER in Epidemiology and Public Health (CIBERESP), Madrid, Spain.

[25]Universidad de Granada, Granada, Spain

[26]Public Health Division of Gipuzkoa, BioDonostia Research Institute, Donostia-San Sebastian, Spain

[27]Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain.

[28]Navarra Public Health Institute, Pamplona, Spain.

[29]IdiSNA, Navarra Institute for Health Research, Pamplona, Spain.

[30]Department of Surgery, Skåne University Hospital Malmö, Lund University, Lund, Sweden

[31]Department of Surgical and Perioperative Sciences, Umeå University, Umeå, Sweden

[32]Institute of Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands

[33]Department of Public Health, the University Medical Center Utrecht, Utrecht, The Netherlands

[34]Medical Research Council (MRC) Epidemiology Unit, University of Cambridge, Cambridge, UK.

[35]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK.

[36]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom

[37]Department of Nutrition, Bjørknes University College, Oslo, Norway

[38]Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway

The entire manuscript is 3,551 words (including abstract (222 words), text (3,329 words), 50 references, 2 tables, 3 supplementary table and 1 supplementary figure).

DISCLAIMER: Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

***CORRESPONDING AUTHOR**: Dr. Raul Zamora-Ros; Unit of Nutrition and Cancer, Catalan Institute of Oncology (ICO), Bellvitge Biomedical Research Institute (IDIBELL), Av Gran Via 199-203, 08908 L'Hospitalet de Llobregat, Spain, E-mail: rzamora@idibell.cat. Telephone number: +34 932607401

AVAILABILITY OF DATA AND MATERIALS: For information on how to submit an application for gaining access to EPIC data and/or biospecimens, please follow the instructions at http://epic.iarc.fr/access/index.php.

FUNDING: This research was funded by the Instituto de Salud Carlos III through the grant CP15/00100 (Co-funded by European Regional Development Fund. ERDF, a way to build Europe). The coordination of EPIC is financially

supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands), the National Institute for Public Health and the Environment (RIVM, Bilthoven, the Netherlands); Health Research Fund (FIS) - Instituto de Salud Carlos III (ISCIII), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology - ICO (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford) (United Kingdom). We thank CERCA Program / Generalitat de Catalunya for institutional support. RZ-R was supported by the "Miguel Servet" program (CP15/00100) from the Institute of Health Carlos III (Co-funded by the European Social Fund (ESF) - ESF investing in your future).

SHORT TITLE: Polyphenol biomarkers & thyroid cancer

KEYWORDS: polyphenol, biomarkers, thyroid cancer, EPIC, nested case–control study

ABBREVIATIONS: EPIC, European Prospective Investigation into Cancer and Nutrition; IARC, International Agency for Research on Cancer; LOQ, limit of quantification; TC, thyroid cancer; TNM, tumor-node-metastasis

ABSTRACT:

Background. Polyphenols are natural compounds with anticarcinogenic properties in cellular and animal models, but epidemiological evidence investigating the associations of these compounds with thyroid cancer (TC) is lacking.

Objective. The aim of this study was to evaluate the relationships between blood concentrations of 36 polyphenols and TC risk in the European Prospective Investigations into Cancer and Nutrition (EPIC).

Methods. A nested case-control study was conducted on 273 female cases (210 papillary, 45 follicular, and 18 not otherwise specified TC tumors) and 512 strictly matched controls. Blood polyphenol levels were analyzed by high pressure liquid chromatography coupled to tandem mass spectrometry after enzymatic hydrolysis.

Results. Using multivariable adjusted conditional logistic regression models, caffeic acid (OR_{log2}=0.55, 95% CI: 0.33, 0.93) and its dehydrogenated metabolite, 3,4-dihydroxyphenylpropionic acid (OR_{log2}=0.84, 95% CI: 0.71, 0.99) were inversely associated with differentiated TC risk. Similar results were observed for papillary TC, but not for follicular TC. Ferulic acid was also inversely associated only with papillary TC (OR_{log2}=0.68, 95% CI: 0.51, 0.91). However, none of these relationships was significant after Bonferroni correction for multiple testing. No association was observed with any of the remaining polyphenols with total differentiated, papillary or follicular TC.

Conclusions. Blood polyphenol levels were mostly not associated with differentiated TC risk in women, although our study raises the possibility that high blood concentrations of caffeic, 3,4-dihydroxyphenylpropionic, and ferulic acids may be related to a lower papillary TC risk.

INTRODUCTION

Thyroid cancer (TC) is the most common endocrine cancer and is classified into two main groups: differentiated (mostly papillary and follicular) and non-differentiated (e.g. anaplastic) carcinomas (1). TC is more frequent in women than in men, and its incidence has been increasing over the last three decades (2), partially attributable to overdiagnosis (3). To date, only few risk factors have been established (i.e. benign thyroid disease, radiation exposure, and body size) (4, 5). However, the role of dietary factors in TC carcinogenesis is not clearly understood (1).

Polyphenols are bioactive phytochemicals, abundant in the human diet and showing a high variability in their chemical structure. Over 500 individual polyphenols have been identified from dietary sources, almost exclusively plant-based foods (6). Once ingested, polyphenols are partially absorbed and conjugated in both the gut mucosa and liver. Many of the non-absorbed compounds reach the colon, undergo extensive catabolism reactions by microbiota, and finally can be absorbed as simple phenolic acids (7, 8).

Established biological properties of polyphenols include antioxidant, anti-inflammatory and chemo-preventive effects (9). Polyphenols have been shown to induce apoptosis, inhibit cell proliferation and invasion in TC cells (10). However, epidemiological evidence on the association between polyphenol intake and TC risk is scarce and inconclusive. In a US cohort, dietary flavan-3-ol intake was negatively and flavanones were positively related to TC risk (11). In a previous analysis of dietary polyphenol intake and differentiated TC risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort the results were

null, except in subjects with BMI≥25, where inverse associations with intake of phenolic acids were detected (12). However, the assessment of polyphenol exposures using dietary questionnaires and food composition databases has well-known limitations. Polyphenol biomarkers constitute an alternative and objective way for estimating polyphenol exposures, which take into account inter-individual variations in bioavailability (13, 14).

We hypothesized that polyphenols may have a preventive role in differentiated TC and polyphenol biomarkers may capture dietary exposure better than questionnaires. Therefore, our aim was to explore the associations between 36 blood polyphenol concentrations and differentiated TC risk, and the difference between TC histological subtypes, in women in a nested case-control study within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort.

MATERIAL AND METHODS

Study population, sample and data collection

The EPIC is an on-going multicenter prospective cohort study that enrolled 521,324 men and women, mainly between the ages of 35 and 70 years, predominantly from the general population of 10 European countries in the nineties (15). All participants gave written informed consent, and the study was approved by the Ethics Review Committee of the International Agency for Research on Cancer (IARC) and by the local ethical committee of individual EPIC centers.

At baseline, habitual food and nutrient intake over the previous year was assessed through a validated center/country-specific dietary questionnaire (15) and the standardized EPIC Nutrient Database (16). Anthropometric data were measured,

except in EPIC-Oxford, Norway and France, where they were self-reported (15). Blood samples, from approximately 80% of the EPIC cohort, were collected at recruitment according to standardized procedures and stored at IARC under liquid-nitrogen (–196ºC) for all countries, except at Denmark under nitrogen-vapor (–150ºC) (15).

Endpoint assessments

Primary incident TC cases were identified through record linkage with regional cancer registries in most of the centers, except in France, Germany, Greece and Naples (Italy), where follow-up was based on a combination of methods, including health insurance records, cancer and pathology registries and active follow-up evaluation of study participants and their next-of-kin. TC was defined as code C73 in the 10th Revision of the International Classification of Diseases (ICD-10). This analysis focused on differentiated TC, i.e., papillary (morphologic codes: 8050, 8130, 8260, 8340–8344 and 8350), follicular carcinomas (8290, 8330–8335), and not otherwise specified, which are likely to also be papillary (8000, 8010, 8140). TC cases with rare or missing histological types (medullary, anaplastic, lymphoma, other morphologies) were not included. For each EPIC center, closure dates of the study period were defined as the latest dates of complete follow-up for both cancer incidence and vital status (between February 2011 and December 2015).

Nested case–control design

Only incident female TC cases were selected among participants with available blood sample at baseline, because the number of TC in men is very low in EPIC (n=76) (17). Female controls were selected by incidence density sampling from all

cohort members alive and free of cancer (except non-melanoma skin cancer) at the time of diagnosis of the corresponding case and were also matched by study center, duration of follow-up, age (± 1 year), date of blood collection (± 3 months), time of blood collection (± 1 hour), and fasting status at the time of blood collection (< 3 hours (not fasting), 3–6 hours (in between), or > 6 hours (fasting)). For every case, 2 matched controls were identified.

Laboratory measurements

Samples, cases and matched controls, were left-overs of a previous EPIC study (17). No samples from Sweden remained for the analysis. Therefore, all samples experienced one freeze-thaw cycle before polyphenol analyses at IARC. Polyphenols concentrations in biological samples are generally stable after freeze-thaw cycles (18, 19). Citrated plasma was used for laboratory analyses except for samples from Denmark (serum). The list of 36 polyphenols measured was tabulated in **Table 1**. Blood polyphenols were measured by differential isotope labeling and liquid chromatography electrospray ionization tandem mass spectrometry. Detailed information of the method was published elsewhere (20). Limits of quantification (LOQ) for the polyphenols varied between 0.11 nmol/L for daidzein and 44.4 nmol/L for quercetin and isorhamnetin. Blood polyphenol concentrations that fell below the LOQ were set to values corresponding to half the limit of quantification. All intra-batch coefficients of variation were <10%; while all inter-batch coefficients of variations were <20% (except for phloretin and enterodiol for which CVs were 22.0% and 21.5%, respectively). Samples from cases and matched controls were analyzed together, within the same analytical batch.

Statistical Analyses

Medians and percentiles (25^{th} and 75^{th}) of blood polyphenol concentrations of cases and controls were calculated and compared using Wilcoxon tests. Spearman's correlation coefficients were calculated to assess the correlations among blood polyphenol concentrations in the controls. Means with standard deviations, medians with interquartile ranges or frequencies (where appropriate) of baseline characteristics were computed and compared between cases and controls. Baseline characteristic differences between cases and controls were tested by conditional logistic regression.

In our power analysis calculations, a total of 273 cases and matched controls (1:2) will allow us to detect an exposure-disease association with a $\beta=0.80$ for an OR=0.6 for highest vs. lowest quartiles of exposure in the control population, assuming $\alpha=0.05$ (21). The estimated disease prevalence is 0.2% (12).

Multivariable conditional logistic regression, stratified by case–control set, was used to compute ORs and the corresponding 95% CI for the associations between blood polyphenol concentrations and differentiated TC risk. The quality of the models was checked using graphical methods and a goodness-of-fit test. Blood polyphenol concentrations were categorized into quartiles based on the distribution of blood levels in controls. Tests for linear trend were performed by assigning the medians of each quartile as scores and entered this variable as a continuous term in the logistic regression models. Blood polyphenol concentrations were also analyzed as continuous variables, after $\log_2$ transformation. OR_{log2} estimates can be interpreted as the relative risk associated with a doubling in blood polyphenol concentration. Possible nonlinear associations were tested using restricted cubic spline models. The basic model was conditioned on matching factors only, while

the multivariable model was additionally adjusted for BMI (kg/m^2), alcohol consumption (g/d), and age of menarche (y). Other lifestyle, anthropometric, and reproductive variables such as smoking status (never, current, former, unknown), physical activity using the Cambridge index (inactive and moderately inactive, moderately active and active, unknown) (22), education level (none, primary, technical/professional, secondary, higher education, unknown), menopausal status [premenopausal, postmenopausal, perimenopausal, surgical postmenopausal (bilateral oophorectomy)], parity (no, yes, unknown), number of full-term pregnancies (nulliparous, 1, 2, 3, ≥4, unknown), breastfeeding (no, yes, unknown), ever oral contraceptives (OC) use (no, yes, unknown), ever hormonal replace therapy (HRT) use (no, yes, unknown), and prevalent diabetes (no, yes, unknown) were evaluated as potential confounders, but were not included in the final model because they were not different (P-value>0.1) between cases and controls in the logistic regressions conditional on matching variables. Missing values were retained by creating a separate category (unknown) for categorical variables.

Similar conditional logistic regression models were conducted for polyphenols (caffeic acid, and 3,4-dihydroxyphenylpropionic acid) which were significantly associated with differentiated TC risk by tumor-node-metastasis (TNM) stage (low: T1-T2 vs. high: T3-T4), and histological type (papillary vs. follicular), and heterogeneity by subgroups was tested using the Wald test assessed with the SAS macro *%subtype* (23). Moreover, modification of the ORs was evaluated by age at blood collection (<48, 48-55, >55 y), education level (primary or lower vs. secondary or higher), smoking status (never vs. ever), physical activity (inactive or moderately inactive vs. moderately active or active), BMI (<25 vs. ≥25 kg/m^2),

menopausal status (premenopausal, perimenopausal, postmenopausal), alcohol consumption (≤5 g/d vs. >5g/d), time to diagnosis (<4, 4-7, >7 years), and countries (high vs. low incidence for differentiated TC) using a likelihood ratio test. EPIC countries with TC incidence rates per year of >1/10,000 in women (i.e., France, Germany, Greece, Italy, and Spain) were considered to have high TC incidence, while UK, the Netherlands, Denmark, and Norway were considered to have low TC incidence.

To account for multiple comparisons, the Bonferroni correction was applied giving a stricter P value threshold for statistical significance at 0.0015, based on the 33 polyphenols analyzed (P value<0.05/33=0.0015). Blood polyphenol levels associated with differentiated TC risk at P value between < 0.05 and 0.0015 were selected as candidates for independent validation studies. All analyses were performed using SAS Software (version 9.3, SAS Institute Inc, Cary, NC, USA).

RESULTS

The current study included 273 incident differentiated TC cases (210 papillary, 45 follicular, and 18 not otherwise specified TC tumors) and 512 matched controls after a median follow-up time of 12.6 years. (**Supplementary Figure 1**). All cases and controls were women with a mean age at blood collection of 50 years. At baseline, controls tended to have a lower BMI and to consume more alcohol than cases (**Table 2**). Moreover, controls were more likely to have experienced menarche at an older age than cases, although the difference was not significant. The rest of baseline characteristics were comparable in cases and controls.

Thirty six polyphenols were measured in blood samples from cases and controls. Three of them (epigallocatechin, gallocatechin, and gallic acid ethyl ester) were excluded from the association analyses because the numbers of samples <LOQ were higher than 75% (Table 1). Most polyphenols showed similar blood concentrations in cases and controls; except caffeic acid found in slightly lower concentrations in differentiated TC cases when compared with controls (Table 1). Moderate correlations were observed between caffeic and ferulic acids (mainly originating from coffee intake) (24) and coffee intake (respectively r=0.39 and r=0.50), and between 3,4-dihydroxyphenylpropionic acid (a metabolite of caffeic acid formed in the gut) and coffee intake (r= 0.38).

Several strong correlations were observed between polyphenol concentrations in blood, such as between 3,5-dihydroxybenzoic acid and 3,5-dihydroxyphenylpropionic acid (r=0.85), genistein and daidzein (r=0.77), naringenin and hesperetin (r=0.72), caffeic acid and 3,4-dihydroxyphenylpropionic acid (r=0.64), and caffeic acid and ferulic acid (r=0.68) reflecting co-occurrence in their main food sources or biotransformation reflecting co-occurrence in their main food sources or biotransformation (**Supplementary Table 2**).

In the multivariable models, blood concentrations of caffeic acid (OR_{log2}=0.55, 95% CI: 0.33, 0.93) and 3,4-dihydroxyphenylpropionic acid (OR_{log2}=0.84, 95% CI: 0.71, 0.99) were inversely associated with differentiated TC risk (**Table 3**), although they did not reach the Bonferroni threshold. In the restricted cubic spline model, no evidence of non-linearity for the relationships between both caffeic acid and 3,4-dihydroxyphenylpropionic acid and differentiated TC risk was observed (data not

shown). All other polyphenol concentrations were not related to differentiated TC risk.

In the results divided by TC histological subtype, inverse associations were observed between blood concentrations of caffeic acid (OR_{log2}=0.36, 95% CI: 0.19, 0.68; P for heterogeneity = 0.048), 3,4-dihydroxyphenylpropionic acid (OR_{log2}=0.74, 95% CI: 0.61, 0.90; P for heterogeneity = 0.030) (**Table 4**), and ferulic acid (OR_{log2}=0.68, 95% CI: 0.51, 0.91; P for heterogeneity = 0.062) and papillary TC tumors; but no associations were detected with follicular TC tumors. None of the other blood polyphenols were associated with either papillary or follicular TC tumors (data not shown). In the subgroup analyses, an inverse association was observed with blood concentrations of caffeic and 3,4-dihydroxyphenylpropionic acids in countries with low TC incidence, but not in countries with high TC incidence (P for heterogeneity < 0.05). However, none of these results reached the Bonferroni threshold (p=0.0015). Similar inverse associations were observed for the relation between either caffeic acid or 3,4-dihydroxyphenylpropionic acid and differentiated TC risk across strata of age at blood collection, education level, smoking status, physical activity, BMI, menopausal status, alcohol intake, and years between blood draw and diagnosis denoting no effect modification (Table 4).

DISCUSSION

In the current prospective nested case-control study, inverse trends were observed between blood concentrations of both caffeic acid and its dihydrogenated metabolite, 3,4-dihydroxyphenylpropionic acid (also called dihydrocaffeic acid) and total differentiated TC risk, but they did not reach the Bonferroni threshold for

statistically significant associations when corrected for multiple comparisons. The remaining blood polyphenol levels were not associated with total differentiated TC risk. Interestingly, the two inverse associations were restricted to papillary TC and were more striking in countries with low incidence of TC. For 3,4-dihydroxyphenylpropionic acid, the negative association was also stronger in stage T1-T2 than in T3-T4 carcinomas. Papillary TC and low stage thyroid tumors are more likely to be related to over-diagnosis than high stage TCs in countries with high incidence. However, over-diagnosis is not related with these TC tumor types in countries with low incidence (3).

To our knowledge, this is the first study evaluating the relations between blood polyphenol levels and TC risk. Although no results were statistically significant after Bonferroni correction, concentrations of caffeic, 3,4-dihydroxyphenylpropionic and ferulic acids might be inversely associated with papillary TC risk, but not with follicular TC risk. Caffeic and ferulic acids are abundant in human diets, and are mostly present in an esterified form as chlorogenic and feruloylquinic acids (esters of caffeic or ferulic acids and quinic acid) (25). They contribute to 78% and 19% of total hydroxycinnamic acid intake (mean intake in Europe = 541.2mg/d) (26). Caffeic acid in blood mainly originates from the hydrolysis of chlorogenic acid by the gut microbiota and from the absorption in the gut of the free form of caffeic acid (27). Ferulic acid in blood results from both the hydrolysis of feruloylquinic acid and from the *O*-methylation of caffeic acid in the liver. Dihydrocaffeic acid is only present in the diet in very low amounts (26). Dihydrocaffeic acid in blood is mainly formed by microbial hydrogenation of caffeic acid in the gut (27). All three compounds in both blood, in the current study, and in urine, in a previous analysis

including 475 subjects from the EPIC study (24), showed moderate-to-high correlations with coffee intake and poor or no correlations with any other tested food groups, except for ferulic acid and cereals (24). Indeed, a urinary metabolite of caffeic acid (caffeic acid sulphate) was correlated to whole-grain rye intake (r=0.58) in a free-living Swedish population (28); while urinary ferulic concentrations were increased after an intervention with rye bran bread in humans (29) and with rye bran in mice (30). Unfortunately, data on coffee consumption was not available in these analyses, so the potential confounding effect of coffee on whole-grain cereal was not measured.

In three previous EPIC studies, intakes of either phenolic acids (mainly hydroxycinnamic acids) (12), or coffee (31), or total fiber (32) were not related to the risk of either overall TC or its histological subtypes (papillary and follicular tumors). Moreover, no differences in the coffee consumption between differentiated TC cases and controls were observed in our study (Supplementary table 1). Furthermore, the consumption of either whole grain cereals or total grains was not associated with TC risk in a series of hospital-based case-control studies (33) or in a meta-analysis (34). Differences between results obtained with the intake measurement, and those obtained here with biomarkers might be explained by a more limited accuracy of exposure measurements when relying on intake data rather than biomarker data (9, 13). In fact, it is difficult to accurately estimate polyphenol intake by dietary questionnaires, due to the variability of polyphenol content within same or similar foods, such as the heterogeneity of polyphenol composition on the different coffee types according to brewing methods (espresso vs. diluted coffee) and cultivars (Arabica vs. Robusta) (35, 36). Thus, dietary

biomarkers should be more accurate and objective measurements than dietary questionnaires, accounting for inter-individual variability in phenolic acid bioavailability (14).

Although the associations were not statistically significant after Bonferroni correction, they were biologically plausible. The underlying potential mechanisms of action of caffeic, ferulic and 3,4-dihydroxyphenylpropionic acids in thyroid carcinogenesis could be directly associated with their anticarcinogenic properties (37). In particular, ferulic acid has been shown to modulate cell cycle arrest, apoptosis, invasion, migration, and colony formation on TT medullary TC cells (38). Moreover, they have been indirectly associated with their anti-diabetic, anti-obesity, indirect antioxidant and anti-inflammatory properties (9). It is important to bear in mind that obesity (5), type 2 diabetes (39) and inflammation (17) are potential risk factors of TC risk. Plasma concentrations of total and several individual polyphenols (i.e. daidzein, 3,5-dihydroxyphenylpropionic acid, 3,4-dihydroxyphenylpropionic acid, ferulic acid, caffeic acid, and hydroxytyrosol) were inversely associated with levels of high-sensitivity C-reactive protein in a previous cross-sectional analysis in an EPIC subsample (40), suggesting that these polyphenols may protect against harmful health effects related to inflammation. Moreover, plasma and urinary concentrations of caffeic acid and other coffee polyphenols were associated with a lower risk of type 2 diabetes in two cohorts (41, 42). Indeed, caffeic and dihydrocaffeic acids inhibit amyloid formation of human islet amyloid polypeptide *in vitro* (43), and decrease glucose uptake and the detrimental effects of high glucose concentrations in endothelial cells (44). In addition, caffeic and ferulic acids modulates the activity of several transcriptional

regulatory factors [e.g. AMP-activated protein kinase (AMPK), peroxisome proliferator activator protein-γ (PPAR-γ), and peroxisome proliferator activator protein-γ co-activator-1α (PGC-1α)] and enzymatic pathways (e.g. fatty acid synthase, 3-hydroxy-3-methylglutaryl CoA reductase, and acyl-CoA cholesterol acyltransferase) to control obesity (45).

Caffeic, ferulic, and 3,4-dihydroxyphenylpropionic acids are compounds of food origin, but they also come from the microbiota catabolism (27). Polyphenols can modulate the gut microbiota towards a more healthy composition (46). Indeed dietary chlorogenic acid supplementation improves gut health in weaned piglets (47). Dysbiosis, an alteration of gut microbiota, is associated with intestinal and extra-intestinal diseases, including cancer and metabolic disorders such as obesity and type 2 diabetes (48, 49). Both TC and thyroid nodules were associated with the composition of gut microbiome in two observational studies in Chinese populations (50, 51).

Major strengths of this study are its prospective design, its long follow-up, its relatively large size for a TC study, and the coverage of several European countries with a wide polyphenol exposure heterogeneity. Moreover, the direct analysis of 36 polyphenols in blood provides a valid measurement of the endogenous exposure. However, several limitations of this study also warrant mention. 1) Half-lives of polyphenols are short to moderate, suggesting that a single measurement of these biomarkers is more likely to reflect relatively short-term levels, except for polyphenols regularly consumed that tend to maintain relatively similar concentrations in blood during the entire day. The three phenolic acids inversely associated with TC risk in the present work mainly originate from

coffee, a beverage most often consumed on a daily basis. 2) Fasting status affects blood levels of polyphenols, particularly polyphenols coming from food and quickly absorbed. However, TC cases were matched with controls by fasting status and time of blood collection to minimize this limitation. 3) We measured blood polyphenols only once for each individual, so we cannot account for intra-variability and changes in the exposure along the study follow-up. This issue could be particularly relevant for few polyphenols, because they have a relatively poor ICC (0.3-0.4), but not for others (ICC >0.5) (http://exposome-explorer.iarc.fr/reproducibilities). Therefore, our results on few blood flavonoids may have been attenuated by partial misclassification. 4) Information on history of benign thyroid diseases, thyroidectomy among control subjects and use of drugs that could interfere with thyroid function was not available in the EPIC study. 5) Although we controlled for a wide range of established TC risk factors, the possibility of residual confounding still exists, though the findings were all little affected by adjustment in our study. 6) We cannot exclude the possibility that our findings were due to chance, because they did not reach the Bonferroni threshold. However, it is often considered to be overly conservative and might have over-corrected the model. Moreover, the findings were similar in both general and subgroup analyses (except for the risk of follicular TC and high TNM stage differentiated TC) and are biologically plausible. 7) Generalizability of the results should be done cautiously, because our study only analyzed European women and other populations may show different genetic background (e.g. non-European ancestry), and microbiota composition with possible consequences on phenolic acid bioavailability.

In summary, this prospective investigation conducted in a relatively large nested case-control study in women within the EPIC, a European multi-country cohort, shows that blood polyphenol concentrations are mostly not associated with TC risk. However, our study raises the possibility that high blood levels of caffeic, 3,4-dihydroxyphenylpropionic and ferulic acids may be related to a lower risk of papillary TC. These three compounds are, therefore, interesting candidates for validation in independent studies on papillary TC.

ACKNOWLEGDEMENTS

We thank Mr Bertrand Hémon for his valuable help with the EPIC database.

CONFLICT OF INTEREST

The authors are not aware of any conflicts of interest.

AUTHOR CONTRIBUTION:

R.Z.-R. designed the research; D.A. performed the laboratory analysis; L.L.-B. performed the statistical analysis; R.Z.-R. drafted the manuscript; S.F., A.S., S.R., A.A. had primary responsibility for final content; S.F., C.K., K.O., A. Tjønneland, T.T., L.L., M.-C.B.-R., V.K., T.S.J., M.B.S., A. Trichopoulou, E.P., C.LV., G.M., V.P., S.P., R.T., F.R., G.S.,JR.Q., M.R.-B., P.A., M.-D. C., E.A., M.A., J.H., R.V., N.J.W., T.Y.N.T., D.A., G.B., E.W., A.S., S.R., A.A. contributed to the design of the study, data collection, and quality control and analysis. All authors read, critically reviewed and approved the final manuscript.

379

REFERENCES

1. Dal Maso L, Bosetti C, La Vecchia C, Franceschi S. Risk factors for thyroid cancer: an epidemiological review focused on nutritional factors. Cancer Causes Control 2009;20:75-86.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
3. Vaccarella S, Franceschi S, Bray F, Wild CP, Plummer M, Dal ML. Worldwide Thyroid-Cancer Epidemic? The Increasing Impact of Overdiagnosis. N Engl J Med 2016;375:614-7.
4. Pellegriti G, Frasca F, Regalbuto C, Squatrito S, Vigneri R. Worldwide increasing incidence of thyroid cancer: update on epidemiology and risk factors. J Cancer Epidemiol 2013;2013:965212.
5. Kitahara CM, McCullough ML, Franceschi S, Rinaldi S, Wolk A, Neta G, Olov AH, Anderson K, Andreotti G, Beane Freeman LE, et al. Anthropometric Factors and Thyroid Cancer Risk by Histological Subtype: Pooled Analysis of 22 Prospective Studies. Thyroid 2016;26:306-18.
6. Perez-Jimenez J, Neveu V, Vos F, Scalbert A. Systematic analysis of the content of 502 polyphenols in 452 foods and beverages: an application of the phenol-explorer database. J Agric Food Chem 2010;58:4959-69.
7. Manach C, Williamson G, Morand C, Scalbert A, Remesy C. Bioavailability and bioefficacy of polyphenols in humans. I. Review of 97 bioavailability studies. Am J Clin Nutr 2005;81(1 Suppl):230S-42S.
8. Williamson G. The role of polyphenols in modern nutrition. Nutr Bull 2017;42:226-35.
9. Zamora-Ros R, Touillaud M, Rothwell JA, Romieu I, Scalbert A. Measuring exposure to the polyphenol metabolome in observational epidemiologic studies: current tools and applications and their limits. Am J Clin Nutr 2014;100:11-26.
10. Shin H-J, Hwang K-A, Choi K-C. Antitumor effect of various phytochemicals on diverse types of thyroid cancers. Nutrients 2019;11:125.
11. Xiao Q, Park Y, Hollenbeck AR, Kitahara CM. Dietary flavonoid intake and thyroid cancer risk in the NIH-AARP diet and health study. Cancer Epidemiol Biomarkers Prev 2014;23:1102-8.
12. Zamora-Ros R, Cayssials V, Franceschi S, Kyrø C, Weiderpass E, Hennings J, Sandström M, Tjønneland A, Olsen A, Overvad K, et al. Polyphenol intake and differentiated thyroid cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Int J Cancer 2020;146:1841-50.
13. Zamora-Ros R, Rabassa M, Llorach R, Gonzalez CA, Andres-Lacueva C. Application of Dietary Phenolic Biomarkers in Epidemiology: Past, Present, and Future. J Agric Food Chem 2012;60:6648-57.
14. Bento-Silva AK, Koistinen VM, Mena P, Bronze MR, Hanhineva K, Sahlstrøm S, Kitrytė V, Moco S, Aura AM. Factors affecting intake,

metabolism and health benefits of phenolic acids: do we understand individual variability? Eur J Nutr 2020;59:1275-93.
15. Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol 1997;26 (Suppl 1):S6-14.
16. Slimani N, Deharveng G, Unwin I, Southgate DA, Vignat J, Skeie G, Salvini S, Parpinel M, Moller A, Ireland J, et al. The EPIC nutrient database project (ENDB): a first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study. Eur J Clin Nutr 2007;61:1037-56.
17. Dossus L, Franceschi S, Biessy C, Navionis AS, Travis RC, Weiderpass E, Scalbert A, Romieu I, Tjonneland A, Olsen A, et al. Adipokines and inflammation markers and risk of differentiated thyroid carcinoma: The EPIC study. Int J Cancer 2018;142:1332-42.
18. Urpi-Sarda M, Zamora-Ros R, Lamuela-Raventos RM, Cherubini A, Jauregui O, de la Torre R, Covas M, Estruch R, Jaeger W, Andres-Lacueva C. HPLC-tandem mass spectrometric method to characterize resveratrol metabolism in humans. Clinical Chemistry 2007;53:292-9.
19. Magiera SB, I.; Kusa, J. Development and validation of UHPLC-ESI-MS/MS method for the determination of selected cardiovascular drugs, polyphenols and their metabolites in human urine. Talanta 2012;89:47-56.
20. Achaintre D, Gicquiau A, Li L, Rinaldi S, Scalbert A. Quantification of 38 dietary polyphenols in plasma by differential isotope labelling and liquid chromatography electrospray ionization tandem mass spectrometry. J Chromatogr A 2018;1558:50-8.
21. Murphy N, Achaintre D, Zamora-Ros R, Jenab M, Boutron-Ruault MC, Carbonnel F, Savoye I, Kaaks R, Kühn T, Boeing H, et al. A prospective evaluation of plasma polyphenol levels and colon cancer risk. Int J Cancer. 2018;143:1620-31.
22. Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, Day NE. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Public Health Nutr 2003;6:407-13.
23. Wang MS, D.; Kuchiba, A.; Lochhead, P.; Kim ,S.; Chan, A.T.; Poole, E.M.; Tamimi, R.; Tworoger, S.S.; Giovannucci, E.; Rosner, B.; Ogino, S. Statistical methods for studying disease subtype heterogeneity. Stat Med 2016;35:782-800.
24. Zamora-Ros R, Achaintre D, Rothwell JA, Rinaldi S, Assi N, Ferrari P, Leitzmann M, Boutron-Ruault MC, Fagherazzi G, Auffret A, et al. Urinary excretions of 34 dietary polyphenols and their associations with lifestyle factors in the EPIC cohort study. Sci Rep 2016;6:26905.
25. Perez-Jimenez J, Fezeu L, Touvier M, Arnault N, Manach C, Hercberg S, Galan P, Scalbert A. Dietary intake of 337 polyphenols in French adults. Am J Clin Nutr 2011;93:1220-8.
26. Zamora-Ros R, Rothwell JA, Scalbert A, Knaze V, Romieu I, Slimani N, Fagherazzi G, Perquier F, Touillaud M, Molina-Montes E, et al. Dietary

intakes and food sources of phenolic acids in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Br J Nutr 2013;110:1500-11.
27. Gonthier MP, Verny MA, Besson C, Remesy C, Scalbert A. Chlorogenic acid bioavailability largely depends on its metabolism by the gut microflora in rats. J Nutr 2003;133:1853-9.
28. Hanhineva K, Brunius C, Andersson A, Marklund M, Juvonen R, Keski-Rahkonen P, Auriola S, Landberg R. Discovery of urinary biomarkers of whole grain rye intake in free-living subjects using nontargeted LC-MS metabolite profiling. Mol Nutr Food Res 2015;59:2315-25.
29. Harder H, Tetens I, Let MB, Meyer AS. Rye bran bread intake elevates urinary excretion of ferulic acid in humans, but does not affect the susceptibility of LDL to oxidation ex vivo. Eur J Nutr 2004;43:230-6.
30. Pekkinen J, Rosa-Sibakov N, Micard V, Keski-Rahkonen P, Lehtonen M, Poutanen K, Mykkänen H, Hanhineva K. Amino acid-derived betaines dominate as urinary markers for rye bran intake in mice fed high-fat diet--A nontargeted metabolomics study. Mol Nutr Food Res 2015;59:1550-62.
31. Zamora-Ros R, Alghamdi MA, Cayssials V, Franceschi S, Almquist M, Hennings J, Sandstrom M, Tsilidis KK, Weiderpass E, Boutron-Ruault M-C, et al. Coffee and tea drinking in relation to the risk of differentiated thyroid carcinoma: results from the European Prospective Investigation into Cancer and Nutrtion (EPIC) study. Eur J Nutr 2019;58:3303-12.
32. Zamora-Ros R, Rinaldi S, Tsilidis KK, Weiderpass E, Boutron-Ruault M-C, Rostgaard-Hansen AL, Tjønneland A, Clavel-Chapelon F, Mesrine S, Katzke V, et al. Energy and macronutrient intake and risk of differentiated thyroid carcinoma in the European Prospective Investigation into Cancer and Nutrition study. Int J Cancer 2016;138:65-73.
33. Liu ZT, Lin A-H. Dietary factors and thyroid cancer risk: a meta-analysis of observational studies. Nutr Cancer 2014;66:1165-78.
34. Akslen LA, Nilssen S, Kvale G. Reproductive factors and risk of thyroid cancer. A prospective study of 63,090 women from Norway. Br J Cancer 1992;65:772-4.
35. Parras P, Martínez-Tome M, Jiménez AM, Murcia MA. Antioxidant capacity of coffees of several origins brewed following three different procedures. Food Chem 2007;102:582-92.
36. Rothwell JA, Loftfield E.; Wedekind R, Freedman N, Kambanis C, Scalbert A, Sinha R. Metabolomic Study of the Variability of the Chemical Composition of Commonly Consumed Coffee Brews. Metabolites 2019;9:17.
37. Rothwell JA, Knaze V, Zamora-Ros R. Polyphenols: dietary assessment and role in the prevention of cancers. Curr Opin Clin Nutr Metab Care 2017;20:512-21.
38. Dodurga Y, Eroğlu C, Seçme M, Elmas L, Avcı ÇB, Şatıroğlu-Tufan NL. Anti-proliferative and anti-invasive effects of ferulic acid in TT medullary thyroid cancer cells interacting with URG4/URGCP. Tumour Biol 2016;37:1933-40.

39. Yeo Y, Ma SH, Hwang Y, Horn-Ross PL, Hsing A, Lee KE, Park YJ, Park DJ, Yoo KY, Park SK. Diabetes mellitus and risk of thyroid cancer: a meta-analysis. PLoS One 2014;9:e98135.
40. Harms LM, Scalbert A, Zamora-Ros R, Rinaldi S, Jenab M, Murphy N, Achaintre D, Tjønneland A, Olsen A, Overvad K, et al. Plasma polyphenols associated with lower high-sensitivity C-reactive protein concentrations: a cross-sectional study within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Br J Nutr 2020:123:198-208.
41. Lee AH, Tan L', Hiramatsu N, Ishisaka A, Alfonso H, Tanaka A, Uemura N, Fujiwara Y, Takechi R. Plasma concentrations of coffee polyphenols and plasma biomarkers of diabetes risk in healthy Japanese women. Nutr Diabetes 2016;6:e212.
42. Sun Q, Wedick NM, Tworoger SS, Pan A, Townsend MK, Cassidy A, Franke AA, Rimm EB, Hu FB, van Dam RM. Urinary Excretion of Select Dietary Polyphenol Metabolites Is Associated with a Lower Risk of Type 2 Diabetes in Proximate but Not Remote Follow-Up in a Prospective Investigation in 2 Cohorts of US Women. J Nutr 2015;145:1280-6.
43. Cheng B, Liu X, Gong H, Huang L, Chen H, Zhang X, Li C, Yang M, Ma B, Jiao L, Zheng L, Huang K. Coffee components inhibit amyloid formation of human islet amyloid polypeptide in vitro: possible link between coffee consumption and diabetes mellitus. J Agric Food Chem 2011;59:13147-55.
44. Natarelli L, Ranaldi G, Leoni G, Roselli M, Guantario B, Comitato R, Ambra R, Cimino F, Speciale A, Virgili F, Canali R. Nanomolar Caffeic Acid Decreases Glucose Uptake and the Effects of High Glucose in Endothelial Cells. PLoS One 2015;10:e0142421.
45. Alam MA, Subhan N, Hossain H, Hossain M, Reza HM, Rahman MM, Ullah MO. Hydroxycinnamic acid derivatives: a potential class of natural compounds for the management of lipid metabolism and obesity. Nutr Metab (Lond) 2016;13:27.
46. Roopchand DE, Carmody RN, Kuhn P, Moskal K, Rojas-Silva P, Turnbaugh PJ, Raskin I. Dietary Polyphenols Promote Growth of the Gut Bacterium Akkermansia muciniphila and Attenuate High-Fat Diet-Induced Metabolic Syndrome. Diabetes 2015;64:2847-58.
47. Zhang Y , Wang Y , Chen D , Yu B , Zheng P , Mao X , Luo Y , Li Y , He J. Dietary chlorogenic acid supplementation affects gut morphology, antioxidant capacity and intestinal selected bacterial populations in weaned piglets. Food Funct 2018;9:4968-78..
48. Marchesi JR, Adams DH, Fava F, Hermes GD, Hirschfield GM, Hold G, Quraishi MN, Kinross J, Smidt H, Tuohy KM, et al. The gut microbiota and host health: a new clinical frontier. Gut 2016;65:330-9.
49. Rinninella E, Raoul P, Cintoni M, Franceschi F, Miggiano GAD, Gasbarrini A, Mele MC. What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. Microorganisms 2019;7:14.
50. Zhang J, Zhang F, Zhao C, Xu Q, Liang C, Yang Y, Wang H, Shang Y, Wang Y, Mu X, et al. Dysbiosis of the gut microbiome is associated with

thyroid cancer and thyroid nodules and correlated with clinical index of thyroid function. Endocrine 2019;64:564-74.

51. Feng J, Zhao F, Sun J, Lin B, Zhao L, Liu Y, Jin Y, Li S, Li A, Wei Y. Alterations in the gut microbiota and metabolite profiles of thyroid carcinoma patients. Int J Cancer 2019;144:2728-45.

Table 1. Medians (25th and 75th percentiles) and number of samples with concentrations below the limit of detection (LOD) of plasma polyphenol levels among differentiated thyroid cancer cases and controls.

Plasma concentrations (nmol/L)	Cases (N=273)		Controls (N=512)		P for differences[2]
	N <LOD	Median (p25-p75)	N <LOD	Median (p25-p75)	
Flavonoids					
Apigenin	1 (0.2%)	10.9 (10.1-12.4)	0	11.2 (10.0-12.7)	0.26
Catechin	112 (41%)	12.0 (5.6-18.2)	215 (41%)	12.2 (5.6-16.9)	0.61
Daidzein	0	7.9 (5.2-17.5)	2 (0.4%)	8.0 (5.6-16.9)	0.61
Epicatechin	132 (48%)	11.4 (5.6-15.4)	292 (55%)	5.6 (5.6-15.0)	0.11
Epigallocatechin[1]	252 (91%)	-	496 (94%)	-	-
Equol	41 (15%)	0.4 (0.2- 0.7)	58 (11%)	0.4 (0.2- 0.7)	0.61
Gallocatechin[1]	274 (99%)	-	523 (99%)	-	-
Genistein	0	4.3 (2.0-11.4)	3 (1.1%)	4.1 (2.2-10.3)	0.97
Hesperetin	68 (25%)	2.3 (1.1-19.3)	142 (27%)	2.2 (1.1- 15.2)	0.67
Kaempferol	0	84.0 (73.0-97.0)	0	84.0 (74.0-94.5)	0.91
Naringenin	8 (1.6%)	3.1 (1.3-11.9)	6 (2.2%)	3.4 (1.6-9.4)	0.84
Phloretin	179 (65%)	1.1 (1.1-2.6)	334 (63%)	1.1 (1.1-2.8)	0.59
Quercetin	0	142.0 (123.0-161.0)	0	142.0 (123.0-165.0)	0.61
Phenolic acids					
3-Hydroxybenzoic acid	2 (0.4%)	17.3 (10.8-30.9)	2 (0.7%)	16.7 (10.9-26.3)	0.53
4-Hydroxybenzoic acid	0	348.0 (313.0-399.0)	0	346.0 (314.5-392.5)	0.71
3,5-Dihydroxybenzoic acid	3 (0.6%)	21.2 (12.3-40.7)	1 (0.4%)	19.1 (11.6-41.3)	0.70
3-Hydroxyphenylacetic acid	17 (3.3%)	53.0 (20.8-101.8)	35 (13%)	56.5 (21.5-108.3)	0.67
4-Hydroxyphenylacetic acid	3 (0.6%)	249.0 (178.0-341.0)	27 (10%)	233.5 (182.0-306.0)	0.22
3,4-Dihydroxyphenylacetic acid	1 (0.2%)	21.8 (16.8-28.4)	2 (0.4%)	21.9 (16.9-28.0)	0.75
3,4-Dihydroxyphenylpropionic acid	13 (2.5%)	18.0 (14.3-26.4)	17 (6.2%)	19.3 (14.6-30.4)	0.053
3,5-Dihydroxyphenylpropionic acid	2 (0.4%)	27.1 (17.0-48.8)	7 (2.6%)	26.5 (17.0-53.5)	0.73

Caffeic acid	0	131.0 (116.0-151.0)	0	135.0 (118.0-157.0)	0.054
m-Coumaric acid	40 (15%)	5.7 (2.6-10.9)	77 (15%)	5.0 (2.1-12.2)	0.63
p-Coumaric acid	0	25.4 (21.2-31.1)	1 (0.4%)	25.6 (21.5- 31.5)	0.50
Ferulic acid	0	104.0 (71.0-183.0)	0	110.5 (71.0-206.5)	0.38
Gallic acid	16 (3.1%)	16.2 (13.7-20.3)	26 (9.5%)	16.1 (13.6-19.9)	0.76
Gallic acid ethyl ester[1]	235 (85%)	-	415 (79%)	-	-
Homovanillic acid	0	82.0 (65.0-106.0)	1 (0.4%)	79.0 (64.0-106.0)	0.59
Isorhamnetin	4 (0.8%)	65.0 (57.0-76.0)	1 (0.4%)	66.0 (57.0-76.0)	0.77
Protocatechuic acid	0	232.0 (215.0-255.0)	2 (0.7%)	230.5 (214.0-257.0)	0.88
Vanillic acid	0	197.0 (178.0-225.0)	2 (0.4%)	195.0 (176.0-230.0)	0.97
Stilbenes					
Resveratrol	106 (38%)	2.5 (1.1-3.9)	199 (38%)	2.5 (1.1- 3.8)	0.91
Lignans					
Enterodiol	62 (23%)	1.0 (0.5-2.0)	110 (21%)	1.0 (0.5- 2.1)	0.55
Enterolactone	4 (0.8%)	8.6 (3.7-15.4)	5 (1.8%)	8.3 (3.8- 15.8)	0.98
Tyrosols					
Hydroxytyrosol	117 (42%)	12.0 (5.6-15.2)	222 (42%)	12.2 (5.6-15.5)	0.37
Tyrosol	0	3.5 (2.7-5.1)	3 (1.1%)	3.7 (2.70-5.3)	0.25

[1]Limit of quantification (LOQ) = 11.1 nmol/L for epigallocatechin and gallocatechin, LOQ = 1.11 nmol/L for gallic acid ethyl ester.

[2]From Wilcoxon tests.

Table 2. Baseline characteristics among differentiated thyroid cancer cases and controls.

Characteristic	Cases (N=273)	Controls (N=512)	P-value[1]
Age at blood collection (y)[2]	50.0 (8.6)	50.0 (8.7)	Matched
Body mass index (kg/m^2)[2]	26.4 (4.7)	25.6 (4.6)	0.007
Alcohol intake (g/d)[3]	1.4 (0.1- 8.1)	2.6 (0.2- 11.2)	0.019
Coffee intake (g/d)[3]	120 (41- 296)	129 (60- 300)	0.82
Age at menarche (y)[2]	12.7 (1.5)	12.9 (1.5)	0.069
Physical activity[4]			0.14
Inactive or Moderately inactive	192 (70.3)	341 (66.6)	
Moderately active or active	80 (29.3)	167 (32.6)	
Smoking status[4]			0.77
Never	162 (59.3)	311 (60.7)	
Former	53 (19.4)	98 (19.1)	
Smoker	53 (19.4)	99 (19.3)	
Highest educational level[4]			0.26
None	28 (10.3)	46 (9.0)	
Primary school completed	98 (35.9)	180 (35.2)	
Technical/professional school	54 (19.8)	86 (16.8)	
Secondary school	38 (13.9)	92 (18.0)	
Longer education	49 (18.0)	103 (20.1)	
Menopausal status[4]			0.47
Premenopausal	128 (46.9)	242 (47.3)	
Postmenopausal	100 (36.6)	194 (37.9)	
Perimenopausal	35 (12.8)	64 (12.5)	
Surgical postmenopause	10 (3.7)	12 (2.3)	
Full term pregnancies (yes)[4]	239 (88.5)	440 (86.4)	0.48
Number of full term pregnancies[4]			0.84
0	31 (11.5)	69 (13.6)	
1	46 (17.1)	85 (16.8)	
2	122 (45.4)	214 (42.2)	
3	48 (17.8)	96 (18.9)	
≥4	22 (8.2)	43 (8.5)	
Breastfeeding (yes)[4]	191 (71.3)	377 (74.8)	0.25
Ever use of OC (yes)[4]	127 (46.5)	242 (47.3)	0.62
Ever use of HRT (yes)[4]	34 (12.8)	69 (13.9)	0.71
Fasting status[4]			Matched
<3h	105 (38.5)	187 (36.5)	
3-6h	41 (15.0)	82 (16.0)	
>6h	125 (45.8)	240 (46.9)	
Prevalent diabetes (yes)[4]	10 (2.1)	5 (1.9)	1.00

OC oral conceptives; HRT hormone replace therapy

[1]From logistic regression conditional on matching variables.

[2]Mean (SD)

[3]Median (25th and 75th percentiles)

[4]N(%).Numbers may not sum to totals due to missing values.

Table 3. Odds ratio (ORs) and 95% confidence intervals (CI) of differentiated thyroid cancer for log_2-transformed polyphenol concentrations (nmol/L).

Polyphenols	Basic model[1]		Multivariable model[2]	
	OR (95% CI)	P-value	OR (95% CI)	P-value
Flavonoids				
Apigenin	0.84 (0.59, 1.20)	0.34	0.83 (0.58, 1.19)	0.32
Catechin	1.06 (0.90, 1.26)	0.47	1.13 (0.95, 1.35)	0.17
Daidzein	0.96 (0.85, 1.09)	0.56	0.96 (0.84, 1.09)	0.54
Epicatechin	1.11 (0.93, 1.33)	0.27	1.13 (0.95, 1.36)	0.17
Epigallocatechin				
Equol	0.95 (0.85 ,1.05)	0.29	0.95 (0.85, 1.05)	0.32
Gallocatechin				
Genistein	1.01 (0.91, 1.11)	0.92	1.00 (0.91, 1.10)	0.98
Hesperetin	1.03 (0.96, 1.09)	0.43	1.02 (0.95, 1.08)	0.62
Kaempferol	1.07 (0.57, 1.98)	0.84	1.05 (0.56; 1.96)	0.89
Naringenin	1.02 (0.95, 1.10)	0.59	1.01 (0.94, 1.10)	0.71
Phloretin	0.96 (0.82, 1.11)	0.56	0.94 (0.81, 1.09)	0.41
Quercetin	0.73 (0.40, 1.35)	0.32	0.81 (0.44, 1.51)	0.51
Phenolic acids				
3-Hydroxybenzoic acid	1.05 (0.90, 1.23)	0.55	1.08 (0.92, 1.27)	0.34
4-Hydroxybenzoic acid	1.24 (0.66, 2.34)	0.50	1.25 (0.65, 2.37)	0.50
3,5-Dihydroxybenzoic acid	0.99 (0.86, 1.14)	0.87	0.99 (0.86, 1.14)	0.88
3-Hydroxyphenylacetic acid	0.99 (0.91, 1.09)	0.91	1.01 (0.92, 1.11)	0.85
4-Hydroxyphenylacetic acid	1.08 (0.86, 1.36)	0.49	1.08 (0.86, 1.36)	0.52
3,4-Dihydroxyphenylacetic acid	0.82 (0.60, 1.10)	0.19	0.83 (0.61, 1.14)	0.25
3,4-Dihydroxyphenylpropionic acid	0.84 (0.71, 0.99)	0.032	0.84 (0.71, 0.99)	0.039
3,5-Dihydroxyphenylpropionic acid	0.99 (0.85, 1.17)	0.94	1.00 (0.85, 1.18)	0.96
Caffeic acid	0.52 (0.31, 0.86)	0.011	0.55 (0.33, 0.93)	0.025
m-Coumaric acid	1.01 (0.93, 1.09)	0.89	1.01 (0.93, 1.10)	0.76
p-Coumaric acid	0.88 (0.62, 1.26)	0.49	0.93 (0.64, 1.34)	0.68
Ferulic acid	0.82 (0.64, 1.04)	0.10	0.82 (0.64, 1.04)	0.10
Gallic acid	0.98 (0.73, 1.32)	0.91	1.06 (0.79, 1.43)	0.71
Gallic acid ethyl ester				
Homovanillic acid	1.02 (0.76, 1.38)	0.88	1.07 (0.79, 1.45)	0.67
Isorhamnetin	0.76 (0.37, 1.57)	0.47	0.71 (0.34, 1.47)	0.36
Protocatechuic acid	0.69 (0.20, 2.40)	0.56	0.76 (0.22, 2.66)	0.66
Vanillic acid	1.05 (0.72, 1.53)	0.81	1.02 (0.70, 1.50)	0.90
Stilbenes				
Resveratrol	0.98 (0.86, 1.11)	0.74	1.03 (0.90, 1.19)	0.63
Lignans				
Enterodiol	0.98 (0.90, 1.08)	0.71	1.00 (0.91, 1.09)	0.93
Enterolactone	0.98(0.89, 1.06)	0.57	0.99 (0.91, 1.09)	0.87
Tyrosols				
Hydroxytyrosol	0.85 (0.67, 1.08)	0.19	0.90 (0.70, 1.14)	0.37
Tyrosol	0.88 (0.71, 1.09)	0.24	0.92 (0.74, 1.14)	0.44

[1]From conditional logistic regressions, conditioned on matching factors only (basic model).

[2]From multivariable conditional logistic regressions, conditioned on matching factors with additional adjustment for BMI, alcohol consumption, and age of menarche.

No associations exceed the Bonferroni threshold ($P<0.05/33$) = 0.0015.

Table 4. Odds ratios (ORs) and 95% confidence intervals (CI) of differentiated thyroid cancer of caffeic acid and 3,4-dihydroxyphenylpropionic acid (log_2-transformed) blood concentrations (nmol/L) stratified by selected variables.

			Caffeic acid		3,4-dihydroxyphenylpropionic acid	
	Cases	**Controls**	**OR (95% CI)**	***P-value***	**OR (95% CI)**	***P-value***
Histological type						
Papillary	210	396	0.36 (0.19, 0.69)	0.048[1]	0.74 (0.61, 0.90)	0.030[1]
Follicular	45	82	1.52 (0.52, 4.49)		1.26 (0.79, 2.01)	
TNM stage						
Low: T1-T2	118	218	0.54 (0.26, 1.10)	0.15[1]	0.73 (0.57, 0.95)	0.040[1]
High: T3-T4	29	56	2.16 (0.46, 10.00)		1.44 (0.90, 2.30)	
Thyroid Cancer Incidence						
High incident countries[3]	219	415	0.86 (0.48, 1.56)	0.010[2]	0.92 (0.77, 1.11)	0.040[2]
Low incident countries[4]	54	97	0.15 (0.04, 0.54)		0.55 (0.34, 0.89)	
Age at blood collection (y)						
<48	114	211	0.43 (0.16, 1.18)	0.57[2]	0.78 (0.58, 1.06)	0.75[2]
48-45	75	135	0.45 (0.13, 1.57)		0.84 (0.59, 1.20)	
≥46	84	166	0.64 (0.31, 1.33)		0.90 (0.69, 1.18)	
Education						
Primary or less	126	226	0.39 (0.14, 1.14)	0.34[2]	0.82 (0.62, 1.10)	0.98[2]
Secondary or more	147	286	0.59 (0.27, 1.31)		0.86 (0.64, 1.17)	
Smoking						
Never	162	311	0.60 (0.24, 1.50)	0.45[2]	0.79 (0.60, 1.04)	0.99[2]
Ever	106	197	0.48 (0.17, 1.32)		1.04 (0.74, 1.47)	
Physical activity						
Inactive or moderately inactive	192	341	0.40 (0.18, 0.89)	0.87[2]	0.71 (0.56, 0.91)	0.31[2]
Moderately active or active	80	167	0.19 (0.03, 1.12)		0.72 (0.41, 1.25)	

Body mass index (kg/m^2)						
<25	119	264	0.56 (0.24, 1.31)	0.28[2]	1.02 (0.78, 1.33)	0.54[2]
≥25	154	248	0.56 (0.24, 1.34)		0.89 (0.69, 1.16)	
Menopausal status at blood collection						
Premenopausal	128	242	0.31 (0.13, 0.78)	0.19[2]	0.78 (0.58, 1.04)	0.60[2]
Perimenopausal	35	64	1.33 (0.16, 10.76)		0.84 (0.52, 1.37)	
Postmenopausal (natural and surgical)	110	206	0.69 (0.35, 1.34)		0.92 (0.72, 1.16)	
Alcohol intake (g/d)						
≤5	176	300	0.90 (0.40, 2.03)	0.23[2]	0.89 (0.70, 1.13)	0.61[2]
>5	96	212	0.42 (0.15, 1.12)		0.83 (0.59, 1.16)	
Years between blood draw and diagnosis						
<4	49	86	1.14 (0.40, 3.31)	0.29[1]	1.02 (0.69, 1.53)	0.17[2]
4-7	56	108	0.38 (0.10, 1.43)		1.01 (0.71, 1.43)	
>7	168	318	0.45 (0.23, 0.89)		0.73 (0.58, 0.91)	

TNM tumour-node-metastasis

[1]P for heterogeneity based on the Wald test.

[2]P for interaction based on the likelihood ratio test.

[3]High incident countries for differentiated thyroid cancer: France, Germany, Greece, Italy, and Spain.

[4]Low incident countries for differentiated thyroid cancer: UK, the Netherlands, Denmark, and Norway.