Food frequency questionnaire is a valid assessment tool of quercetin and kaempferol intake in Iranian breast cancer patients according to plasma biomarkers

Sara Sadeghi [a, b, c], Vahid Montazeri [d], Raul Zamora-Ros [e], Pourya Biparva[c], Siamak Sabour [f], Saeed Pirouzpanah [a*, b]

[a] Molecular Medicine Research Center, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran. SP: pirouzpanah@gmail.com, SS: sarasadeghi.09@gmail.com,

[b] Department of Biochemistry and Dietetics, Faculty of Nutrition and Food Sciences, Tabriz University of Medical Sciences, Tabriz, Iran.

[c] Department of Nanotechnology, Islamic Azad University, Langaroud Branch, Langaroud, Iran. p.biparva@sanru.ac.ir

[d] Department of Thoracic Surgery, Faculty of Medicine, Tabriz University of Medical Sciences/ and also Surgery Ward, Nour-Nejat Hospital, Tabriz, Iran. vahid.montazeri1950@gmail.com

[e] Unit of Nutrition and Cancer, Cancer Epidemiology Research Programme, Catalan Institute of Oncology (ICO), Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain. rzamora@idibell.cat

[f] Department of Clinical Epidemiology, Faculty of Health and Safety/ and also Safety Promotion and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, Tehran 198353-5511, Iran. s.sabour@sbmu.ac.ir

*** Corresponding author:**

Saeed Pirouzpanah, M.Sc., Ph.D.
Molecular Medicine Research Center,
Tabriz University of Medical Sciences,
Tabriz 5166614711, Iran.
Tel: +989123490037 (mobile phone)
+9841-33387580-3 (310) (office)
Fax: +9841-33368805
E-mail: pirouzpanah@gmail.com
pirouzpanahs@tbzmed.ac.ir

Abbreviations

ALNM; axillary lymph node metastasis

AUC; area under the curve

B; type II error

BMI; body mass index

BCRL; Breast Cancer Risk and Lifestyle

CI; confidence interval

FFQ; food frequency questionnaire

HPLC; high performance liquid chromatography

Kcal; kilocalorie

LOD; limits of detection

LOQ; limit of quantification

N; number

r, Pearson's correlation coefficients

rs; Spearman's rho correlation coefficients

ROC; receiver operating characteristic

SD; standard deviation

STROBE-Nut; the strengthening the reporting of observational studies in epidemiology-nutrition

SE; standard error

SR; standard reference

USDA; United States Department of Agriculture

Abstract

Our study aimed to assess the accuracy of dietary flavonol (quercetin, kaempferol, and isorhamnetin) intake from a food frequency questionnaire (FFQ) compared to fasting plasma flavonol concentrations, as biomarkers of exposure, in breast cancer patients. In a consecutive case series, newly diagnosed patients with breast cancer (n=140) were recruited at Nour-Nejat Hospital, Tabriz, Iran. Flavonol intake was assessed using a validated FFQ. Plasma flavonol levels were measured using high-performance liquid chromatography-ultraviolet detector. The accuracy of dietary status was evaluated using a receiver operating characteristic (ROC) and area under the ROC curve (AUC). Dietary status was shown in dichotomous using ROC-cutoff point.

The plasma concentrations of quercetin were moderately correlated with dietary intake of quercetin (Spearman's correlation coefficient (r_s) =0.188, P<0.05; $r_{partial}$=0.330, P<0.01) and plasma concentrations of isorhamnetin (r_s =0.337, P<0.001). A linear correlation between dietary levels and plasma concentrations of kaempferol was attained ($r_{partial}$=0.240, P<0.05). Using a ROC-cutoff of 61.9 nmol/L for plasma quercetin (test reference), we were able to differentiate between lower and higher consumers of quercetin with an $AUC_{ROC\text{-}based\ reference}$ =0.65 (P<0.01, sensitivity=61.8%, and specificity=60.0%). Using a plasma kaempferol level of 60.1 nmol/L (ROC-cutoff), it was possible to detect significant differences between higher and lower intakes of kaempferol ($AUC_{ROC\text{-}based\ reference}$ = 0.64, P<0.05).

The correlations and diagnostic performance with plasma concentrations could present a significant accuracy rate (validity), which seems acceptable for a nutritional questionnaire (FFQ) to assess quercetin and kaempferol. An improvement in the classification accuracy of flavonol exposure can provide more precise chemopreventive effects of flavonols in humans, increasing their clinical significance.

***Keywords*:** Test accuracy; Food frequency questionnaire; Biomarker; Quercetin; Kaempferol; Isorhamnetin.

Introduction

Flavonoids comprise a class of plant polyphenolic compounds that are ubiquitously found in many fruits and vegetables [1, 2]. Flavonol is a subclass of flavonoids and a major constituent of the human diet [2]. The configuration of multiple hydroxyl groups in a flavone structure forms a diversity of flavonols [3, 4]. They are well-known for their biological properties, such as antioxidant, anti-mutagenic, anti-proliferative, and pro-apoptotic effects, as well as cell transduction regulation [5]. Quercetin (3,5,7,3',4'-pentahydroxylflavone) and kaempferol (3,5,7,4'-tetrahydroxylflavone) are the major flavonols consumed in the human diet [6]. Both urinary and plasma concentrations of flavonols have been used to measure the absorption and excretion of flavonols in pharmacokinetic and dose-response studies [4, 7]. Measuring urine biomarkers of flavonoids is a sensitive method when normalized in concentration to urinary creatinine, easier to measure in the laboratory, and suggested to have a longer half-life [8]. However, most of the large epidemiological studies only collected blood samples [9-12]. Indeed, biomarkers are expected to represent cellular exposure [13], while, in addition, the estimated area under curve (AUC) values of urinary flavonoids are usually less than plasma biomarkers [14]. Although quercetin in plasma is sufficiently sensitive to represent low intakes of quercetin-rich food, uncertainty remained about considering the corresponding metabolites as biomarkers [15]. Plasma levels of isorhamnetin, the 3′-O-methoxylated derivative of quercetin, could potently serve as a secondary biomarker to co-evaluate quercetin intake status. Studies have documented a significant correlation between plasma concentrations of isorhamnetin and quercetin [12]. Many studies have recorded plasma concentrations of kaempferol (aglycon) in correlation with the intake status of this flavonoid [15-17]. Plasma kaempferol (aglycon) reflects more truly the intake status because it excreted unmetabolized after absorption [15, 18]. Moreover, DuPont et al. found out higher concentrations of free kaempferol (aglycon) in plasma (40% of total kaempferol) than urinary

132 measures (16% of total kaempferol), implicated by the higher affinity of glucuronidase to
133 hydrolyze kaempferol-3-glucuronide and high kidney metabolism in pre-excretion of
134 kaempferol [15, 17].

135 Although *in vitro* studies suggest anti-carcinogenic effects of flavonols in a wide range of
136 cancer cell types, results in epidemiologic studies are inconsistent [19]. This could be in part
137 due to limitations of the dietary assessment using questionnaires [20], which are usually
138 prone to random and systematic errors [21-23]. Moreover, plasma and urinary concentrations
139 of flavonoids are usually affected by intra- and inter-individual variations related to
140 absorption, intestinal-hepatic metabolism, and excretion, and therefore, these could add
141 difficulties to the selection of a precise biomarker or led to controversies in epidemiological
142 studies *[24-26]*.

143 The food frequency questionnaire (FFQ) is the most commonly used nutritional tool to
144 evaluate dietary intake in epidemiological studies [27-29]. It is essential to bear in mind that
145 the validity of FFQ-based data depends on the degree to which FFQ can truly measure dietary
146 intake [22].

147 The validity of the FFQ is usually performed by correlating to another type of
148 questionnaire (such as 24-hour recalls or records) [30]. A popular method of validity applied
149 in epidemiology is the triad method [30]. However, assimilating two sets of questionnaires in
150 the method of triads is nearly always associated with the overestimation of regression
151 coefficients because of similarity in the source of errors or erroneous in both cases [24, 31].
152 Instead, measuring surrogate dietary biomarkers is independent on most biases associated
153 with dietary assessment techniques when used as reference [32, 33].

154 The plasma half-life of flavonols has been documented to be below eight hours, but it
155 could be extended to a day when considering microbiota-derived polyphenolic metabolites
156 [34, 35] or continue the consumption of dietary sources flavonols [15].

Iranian dietary habits are similar to the Mediterranean diet and consist of a variety of plant-derived foods and indigenous vegetables, such as fruits, *Allium* vegetables (rich in quercetin), saffron, and leafy green vegetables (rich in kaempferol) [36]. Thus, an accurate measurement of flavonol exposure is needed to investigate the contribution of flavonols to the health outcomes of Iranian populations. The present study aimed to evaluate if an FFQ is a valid instrument to assess dietary intake of quercetin, kaempferol, and isorhamnetin compared to plasma biomarkers, as references, in a population of newly diagnosed breast cancer patients. Results of this study can improve the way to estimate the dietary exposure of flavonols truly and accurately investigate their relationships with health outcomes, especially cancer, in epidemiological studies.

Methods and Materials

Study design

The Breast Cancer Risk and Lifestyle (BCRL) study is a prospective large consecutive case series on breast cancer patients who were histopathologically diagnosed with primary breast cancer. It is a multicenter study designed to assess lifestyle-related factors in association with breast cancer risk prevention, regional to northwestern Iran, and began in May 2009. The current study is a part of this cohort case series, which included eligible 140 cases recruited and interviewed before surgery (modified radical mastectomy or breast-conserving surgery) at Nour-Nejat Hospital and Shahryar Hospital (Tabriz, Iran). The recruitment period for this analysis was from February 2012 until June 2014.

Participants

The inclusion criteria were the following: women newly diagnosed with invasive ductal breast carcinoma, frequently from the histological grade of 2 or 3, who were willing to participate and signed a consent form before inclusion in the study. The exclusion criteria because they can disturb homogeneity of the study population [37], influence the metabolism

or intake levels of flavonols were considered as follow: the history of previously confirmed malignancy, no history of local or distant metastasis, suffering from complicated disorders that can influence the metabolism of flavonols (liver or kidney dysfunction), gastrointestinal inflammatory disorders (such as gastritis, peptic ulcer, and inflammatory bowel syndrome)[38, 39], being previously exposed to adjuvant therapies, long-term drug use (tamoxifen, raloxifene, methotrexate, theophylline, metformin, anticonvulsants, cyclosporine, epilepsy-related drugs, contraceptives, and hormone replacement therapy) [38, 40, 41], being pregnant [40], postpartum or breastfeeding at the time of diagnosis [40], following a vegetarian diet, taking alternative medicines (such as homeopathy), and body mass index (BMI) >40 kg/m^2 [37].

Three hundred and eighty-four women with breast cancer (2012-2014) were invited to the study of whom 282 agreed to participate (participation rate =73.4%). After exclusions, 170 women were selected of whom 30 were also excluded for histopathological disinformation (diagnosis by surgical dissected tissue ruled out the pre-surgery results by fine-needle aspiration) or insufficient collection of blood samples. Finally, 140 subjects were included in this analysis (Fig. 1). This report was prepared following the Strengthening the Reporting of Observational Studies in Epidemiology—Nutritional Epidemiology (STROBE-Nut) statement specified for nutritional epidemiologic studies [42], and details were listed in Supplementary Table 1.

Participants with a ratio of total energy intake to basal metabolic rate <1.14 were classed as under-reporters for energy according to cut-offs proposed by Goldberg et al. [43]. Basal metabolic rate was calculated using Schofield equations [44].

Ethics approval and consent to participate

All participants signed an informed consent form prior to enrollment. The research protocol outlining the methodology, study subjects, sample size, data collection, biochemical tests and

analysis, and related ethical considerations have been reviewed and received ethical approval by the Ethical Committee of Tabriz University of Medical Sciences (Ethical code: 5-4-1699).

Dietary assessment

The Block's Health Habit and History Questionnaire (FFQ) [14] was translated into Farsi (Persian) using the standard method of "backward-forward" to develop a modified FFQ. Details of the content and face validity are described in Supplementary Materials 1. A trained clinician carried out face-to-face interviews with each participant. The final semi-quantitative FFQ includes 136 food items, 25 questions on food preparation, and 25 fields for open-ended questions. The FFQ was previously validated for folate and cobalamin, using biomarkers, in a different population of women who were newly diagnosed with breast cancer in Tehran [37, 45-47] and recently in Tabriz [48]. Briefly, the FFQ was divided into ten specified food groups, including bread and cereals, dairy products, meat, legumes, nuts, fruits, vegetables, oil, beverages, and spices. The subjects were asked about the average frequency of intake of each food during the previous year before diagnosis. The frequency of each consumed food item was asked on a daily, weekly, monthly, and yearly basis, with the additional response option of never. The standard portion size was defined for participants according to serving sizes or familiar household units such as cup, tablespoon, slice, patty, link, and others. A set of photographs and usual household measurements were also used to help participants to accurately recall portion sizes.

The composition of quercetin, kaempferol, and isorhamnetin in foods (mg aglycone equivalents/100 g) was obtained from the United States Department of Agriculture (USDA) database on flavonoid , release 3.2 [49]. Cooked data from the USDA database was selected when it was available. Nutritionist IV software (Version 3.5.2; 1994, N-squared Computing, SanBruno, CA, USA) was used to calculate the total energy intake of each participant.

Biomarker assessment

The standards for quercetin (purity≥95%; Product No: Q4951), kaempferol (purity ≥97%; Product No: 60010), and isorhamnetin (purity≥95%; Product No: 17794) were obtained from Sigma-Aldrich (St. Louis, MO, USA). Beta-glucuronidase/sulfatase was also purchased from Sigma-Aldrich (Cat. No.: G-7017; St. Louis, MO, USA). All reagents used were high-performance liquid chromatography (HPLC) grade.

Fasting venous blood samples were collected in tubes containing K3-ethylenediaminetetraacetic acid (Sunphoria, Taipei, Taiwan), centrifuged at 1370 g for 15 min, and plasma samples were stored at -70°C until analysis. The method used to hydrolyze and extract flavonols was previously described [9, 50]. Plasma samples (250 μL) were acidified with 50 μL acetic acid (0.5 M/L) and added 100 μL naproxen (25 μg/mL of methanol) as an internal standard. After adding 22 μL (≥220 units) of beta-glucuronidase/sulfatase from Helix Pomatia (Type HP-2, aqueous solution, ≥100,000 units/mL, Sigma-Aldrich; St. Louis, MO, USA), the mixture was vortexed for 1 min and then incubated for 30 min at 37°C. Flavonols (quercetin, kaempferol, and isorhamnetin) were extracted with acetone (six-fold of the sample volume) by shaking the microtubes for 20 min. After centrifuging at 8000 g (10 min, 4°C), the supernatant was evaporated under nitrogen. The residue was resolved in 400 μL methanol [9, 50]. Plasma flavonol concentrations were measured using HPLC (Waters 1525, Binary HPLC Pump, Waters 717 Plus Auto-sampler and Waters 486, Tunable Absorbance Detector, Milford, MA, USA). For HPLC analysis, the final solution of 100 μL was injected into a C18 column (4.6×250 mm, particle size: 4 μM; Waters Nova-Pack, USA) and maintained at 25°C. The mobile phase was composed of a methanol/water mixture (60/40) and 0.2% phosphoric acid. The flow rate was 1 mL/min. Detection was carried out using an ultraviolet detector (Waters 486, Tunable Absorbance Detector, USA). Flavonols were detected at 370 nm (Supplement

Fig. 1a). The acquisition and processing of chromatography data were evaluated using the Waters' Millennium software (Ver. 32). Plasma flavonols were identified by comparison with the retention time of the individual standards (Supplement Fig. 1b). A standard calibration curve was created for each flavonol (Supplement Fig. 1c). Limits of detection (LOD) and limit of quantification (LOQ) of quercetin were 0.0234 ng/ml (0.258 nmol/L) and 0.0782 ng/ml (0.258 nmol/L), respectively. The LOD and LOQ were quantified for kaempferol in 0.0233 ng/ml (0.0814 nmol/L) and 0.0778 ng/ml (0.2718 nmol/L), and for isorhamnetin in 0.039 ng/ml (0.1233 nmol/L) and 0.132 ng/ml (0.417 nmol/L), respectively. The least concentration of the calibration curve extrapolated was not below 10 ng/ml, which is higher than all LOQs [51, 52]. The within-subject coefficient variation was estimated at 4.87% for quercetin, 4.31% for kaempferol, and 4.48% for isorhamnetin. Each plasma sample was tagged with a specific number to make laboratory operators blinded to the sampling data.

Statistical analysis

The sample size was calculated by taking into account the comparisons of means between the highest vs. the lowest categories (quintiles) or consumers vs. non-consumers of a specific diet, food, or beverage (Supplementary Table 2). The power of analysis was assumed at 80% [1- β (type II error)] and the level of significance at 95% [1-α (type I error)=0.05] to ensure that the specified precision involved in the calculation of needed sample size. The expected variability of measurements using different sources of data expressing total flavonoids, total flavonols, and certain flavonol (quercetin, kaempferol, isorhamnetin, and isoflavonoid) of studies [mean±standard deviations (SD)] listed in the Supplementary Table 2. On average, the approximate number of patients per group (low or high intake of “total flavonols”) was 67.2 or rounded up to 140 participants in all.

Data distribution of both dietary intake levels and plasma concentrations of flavonols was evaluated by the Kolmogorov-Smirnov test. Scatter plots were created, and Spearman's

correlation coefficients (r_s) were used to the relation between dietary intake levels and the corresponding plasma concentrations of quercetin, kaempferol, and isorhamnetin (continuous variables). The equation of linear regression was individually considered for each panel of scatter plots accompanying the related standard error for the estimated slope coefficients of linear regression. Partial correlation generated r adjusted for body mass index (kg/m^2), age at first mense (years; y), age at first pregnancy (y), and the number of breastfed child(ren). The receiver operating characteristic (ROC) curve was generated based on plotting the sensitivity of a test (Y-axis) against the false-positive error rate (known as 1-specificity shown in the X-axis) for models with binary classification. The objective of test research was to assess whether a single accuracy test (index test) adequately can show the presence or absence of a particular condition, which was defined as high intake or low intake status [53]. The area under the ROC curve (AUC) was an outcome measure to identify the test accuracy in interpreting the ability of dietary flavonols (test variable) to correctly differentiate between high vs. low plasma concentrations (reference variables or predictor). Plasma concentrations of biomarkers (quercetin, kaempferol, and isorhamnetin) were dichotomized (high vs. low concentrations) with different cutoffs as follows: 1) cutoffs provided by a previous study or published values nominated as standard reference (SR) (quercetin ≥80.2 nmol/L and kaempferol ≥57.8 nmol/L) [9]; 2) receiver operating characteristic (ROC)-based cutoffs defined by Youden's index and explored ROC-based reference which are close to the values in former publications (quercetin ≥61.9 nmol/L [10, 54] and kaempferol ≥60.1 nmol/L [9]); 3) median as a cutoff determined at present sample population was identified in interpreting the median model (quercetin ≥85.9 nmol/L and kaempferol ≥67.6 nmol/L). Studies showed menopausal status as a hormone-related effect modifier potently assigned to quercetin metabolism [40], and accordingly, menopause, when HRT was an exclusion criterion, was considered a potential confounder [41].

Energy-adjusted flavonols were calculated according to the residual method described by Willett [32]. All statistical tests were two-tailed, and the statistical significance level was considered at $P<0.05$. Data were analyzed using SPSS 16.0 (SPSS Inc., Chicago, IL, USA).

Results

Participants

One hundred and forty women were included in the final analysis. The mean age at the diagnosis of participants was 46.3 years (SD=8.5). The mean BMI was 28.6 kg/m² (SD=4.4). Premenopausal women were more common (85 out of 121, 70.2%) than postmenopausal patients ($P<0.001$). The frequencies of histopathological grades included were 15.7%, 71.9%, and 12.4% for grade I, grade II, and grade III, respectively.

General and clinicopathological features and flavonol levels

Table 1 summarizes participants' characteristics (general information, pathologic data, and anthropometric indices) by the binary status of plasma flavonols, but no significant result could support the observed differences.

Supplementary Fig. 2 illustrates the plasma concentrations of studied flavonols between the dichotomous status of the hormonal receptor [estrogen receptor (ER) and progesterone receptor (PR)], human epidermal growth factor receptor (HER-2), and among the molecular subtypes (luminal sub-classes) diagnosed in breast tumors. The plasma concentrations of kaempferol was observed higher among patients with ER-negative tumors ($P<0.05$). However, both quercetin and kaempferol showed higher plasma concentrations when patients diagnosed with PR-negative tumors ($P<0.05$). Plasma concentrations of kaempferol were significantly higher in patients with triple-negative tumors than the luminal A sub-class ($P<0.05$).

Intake levels and plasma concentrations of flavonols

Dietary intake levels and plasma concentrations of flavonols (means and percentiles) are presented in Table 2. The highest mean intake of flavonols was for quercetin (67.6 mg/day) followed by kaempferol (24.4 mg/day), and by isorhamnetin (4.3 mg/day). The mean plasma concentrations of quercetin, kaempferol, and isorhamnetin (including other methylated metabolites of quercetin) were 102.5 nmol/L, 74.3 nmol/L, and 897 nmol/L, respectively.

Table 3 presents the correlation coefficient values between the number of servings of food groups and the plasma concentration of flavonols. Plasma concentrations of quercetin were correlated with the intake amounts of fruits (correlation coefficient (r) =0.228, $P<0.01$) and legumes (r=0.176, $P<0.05$). Plasma kaempferol showed a significant correlation with vegetable intake (r=0.241, $P<0.01$).

Among all investigated food items (n=136), there were only statistically significant correlations between raisin (r=0.209), nectarine (r=0.180), grapes (r=0.180), chocolate (rs=0.215), leek (r=0.234), cabbage (r=0.219), onion (r=0.235), spring onion (r=0.235), spinach (r=0.216), ginger (r=0.190), cucumber (r= 0.187), red beans (r=0.181) and pistachio (r=0.197) and plasma levels of quercetin. Correlation were also found between banana consumption, cabbage and saffron with plasma levels of kaempferol with r coefficient ranging from 0.170-0.228.

Diagnostic or test's performance

Scatter plots [Spearman's correlation coefficient (r_s)] are shown in Fig. 2 to determine whether dietary intake levels of quercetin and kaempferol (test variables) are associated with the corresponding plasma concentrations of biomarkers (reference variable). The daily dietary intake of both quercetin in mg/day (r_s =0.17, $P<0.05$) and energy-adjusted quercetin (r_s =0.19, $P<0.05$) were significantly correlated with plasma concentrations of quercetin. After adjustments for potential covariates, the effect size of the correlation between dietary

intake levels and plasma concentrations of quercetin was almost above moderate (quercetin: $r_{partial.}$=0.33, residual quercetin: $r_{partial.}$=0.31, P<0.01). The plasma level of isorhamnetin was significantly correlated with the plasma level of quercetin (r_s =0.34, P<0.001), but not with dietary isorhamnetin. Unlike dietary intake of kaempferol did not show any correlations with plasma concentrations of kaempferol (r_s =0.09, P=0.27), adjustment for potential covariates showed weak-to-moderate significant correlations between dietary intake levels and plasma concentrations of kaempferol (kaempferol: $r_{partial.}$=0.245 and residual kaempferol: $r_{partial.}$=0.230, P<0.05).

Cross-tabulation was carried out to evaluate the association of dietary status of flavonoids (high and low) and the dichotomous status of plasma flavonoids (high and low levels) as a reference biomarker (Supplementary Table 3). High dietary intakes of quercetin were frequently reported in the group of subjects who had high plasma concentrations of quercetin (P<0.05, Supplementary Table 3). When the tests were repeated for premenopausal women, all of the models remained consistently significant (P<0.05). Among premenopausal women, high kaempferol intake was often classified as a high plasma state of kaempferol and deemed significant in all of the models (P<0.05; Supplementary Table 3).

Supplementary Table 4 showed that the intake levels of quercetin from cooked vegetables had positive correlation with plasma concentrations of quercetin ($r_{partial}$=0.217, P<0.05), while raw vegetables did not show such correlation. Dietary intake of kaempferol from raw vegetables could present significant correlations with plasma concentrations of kaempferol ($r_{partial}$ =0.301, P<0.01). Dietary quercetin had significant correlations with intake levels of raw vegetable (quercetin: r_s=0.261, P<0.05), cooked vegetables (r_s=0.573, P<0.001) and fruits (raw, r_s=0.269, P<0.001). Dietary kaempferol showed strong correlation with raw vegetables (r_s=0.434, P<0.001) and significant link with fruit intake (r_s=0.198, P<0.05, Supplementary Table 4).

To quantify the distinguishing ability of ROC-, the median- or SR-based cutoff to classify plasma concentrations of flavonols, the related test accuracy was evaluated (Table 4). Quercetin FFQ data was able to significantly differentiate between high and low plasma concentrations of quercetin using the ROC-based reference model (AUC=0.65; P=0.004), the SR model (AUC=0.61; P=0.029), and the median model (AUC=0.60; P=0.044) (Table 4). Among premenopausal women, similar findings were indicating the association between the dietary intake status of quercetin and biomarkers in the ROC-based reference model (AUC=0.69, P=0.004), SR model (AUC=0.65, P=0.021), and median model (AUC=0.63, P=0.040).

In premenopausal women, the test accuracy of dietary kaempferol was attained in the SR model at AUC=0.66 (P=0.066) and the ROC-based reference model at AUC=0.64 (P=0.044, sensitivity= 58.7% and specificity= 72.7%) (Table 4).

Discussion

Our findings showed that the intake assessments of quercetin and kaempferol by a FFQ had a significant accuracy rate (validity) for detecting actual status based on biomarker measures in breast cancer patients. There was a moderate linear correlation between dietary intake levels and plasma concentrations of quercetin, the weak-to-moderate correlations between measures of kaempferol, which demonstrated the agreement between nutritional data and a biomarker concentration. The pairwise correlation between plasma concentrations of isorhamnetin and quercetin was also notable, supporting quercetin as an accurate biomarker.

Many biomarker-based validity assays are supporting the use of dietary questionnaires for the assessment of dietary flavonoids [16, 55]. Consistent with the present findings showed significant linear correlations between quercetin (crude and energy-adjusted dietary intakes) and plasma concentrations of quercetin, Zhang et al. [21] reported a significant agreement

406 between the mean intake levels of quercetin, kaempferol, and isorhamnetin (FFQ data) and
407 plasma concentrations of surrogate biomarkers in Chinese university campus population.
408 Radtke et al. [16] have shown that the habitual dietary intake of flavonols was correlated with
409 fasting plasma concentrations of quercetin and kaempferol. They suggested that fasting
410 plasma concentrations of flavonoids are valid nutritional biomarkers [16, 21], reflecting the
411 dietary intake of flavonoids, especially over a short time [9, 50, 56]. Noroozi et al. [57]
412 studied the correlation between high levels of flavonoid intake in the habitual diet of a
413 population who already had been supplemented by flavonols and demonstrated remarkable
414 correlations with plasma and urine concentrations of quercetin. Supplementation studies to
415 assess flavonoid intake could reasonably provide greater correlation coefficients [1, 54].
416 Correlation coefficients have varied between weak and strong in many studies [9, 10, 16].
417 Apart from dietary vs. pharmacological doses [1, 54], differences in correlations might be
418 partially due to the use of different databases for flavonoid contents (taking into account
419 missing values and variability in the composition of these compounds in foods), and intra-
420 and inter-variability in the pharmacokinetics of flavonols [1, 35, 57]. More importantly,
421 short-term dietary assessments (such as 24h dietary-recall) are methods usually accompanied
422 by stronger correlations than long-term assessments such as FFQ) [25, 58]. The included food
423 groups are not uniform across studies, and therefore, this might lead to variation in
424 correlations. Moreover, studies had only a single measurement of flavonoids that would
425 underestimate the validity of habitual exposure [59].

426 Plasma concentrations of quercetin had a significant correlation with plasma
427 concentrations of isorhamnetin in terms of measuring methylated derivatives of dietary
428 quercetin with acceptable intra-assay precision. Therefore, the considerable collinearity
429 between both biomarkers supports the accuracy of plasma concentrations of quercetin to be

430 considered as a reference biomarker [12]. However, studies have been unsuccessful in fully
431 validating plasma concentrations of kaempferol as biomarkers of dietary kaempferol [12, 60].
432 Apart from several studies suggesting the validity of FFQ in the assessments of
433 phytoestrogen intake of Asian [61-63], and even other populations [64, 65], limited
434 information put efforts on flavonol intakes validity [21, 66], particularly based on
435 biomarkers. In addition to *Allium* (onion and spring onion), cruciferous and
436 Apiaceous (vegetables) and fruit intakes correlated highly with querctin and kaemferol
437 intakes, studies in Asian population showed the major source of flavonols (and flavones) in
438 fruits (apple and citrus) and vegetables (potato and celery) [66, 67], and Western population
439 suggested beverages as the sources of dietary polyphenols [30]. This suggested the variety of
440 food composition in FFQ world-wide which is concerning to the geographical variations
441 including typical habitual diets in populations.
442 One of the most important validity parameters that can address the diagnostic (test)
443 accuracy of classification is AUC [45, 68]. According to AUC-based findings, the use of FFQ
444 data (test indicator) was shown to have an adequate ability to accurately detect the quercetin
445 intake status attained in both SR- and median-based models. The test's accuracy on quercetin
446 supports the linear correlation between dietary and plasma concentrations and strengthens the
447 results of dietary accuracy of kaempferol, particularly in premenopausal women. It is also
448 possible that menopause, a state of lacking estrogen, may partly mediate the association
449 between diet and plasma flavonols in breast cancer [41]. Thus, one contribution of the current
450 analysis is that FFQ could be considered a valid dietary assessment tool to provide a suitable
451 method for the classification of quercetin intakes and somewhat an acceptable kaempferol
452 tool intake [69]. In addition to the overriding spells of quantitative data in validation studies,
453 for future studies, validating the stratifying criteria could improve our understanding and,
454 therefore, help us to interpret more precisely the diet-related cancer risk.

Although *in vitro* and experimental animal studies showed that quercetin and kaempferol could suppress proliferation and regulate cell transduction [5]; epidemiological data are not conclusive to support anti-proliferative effects of flavonols, especially in breast cancer [70]. The present FFQ was generated and designed to assess the dietary intake of nutrients and non-nutritive bioactive components (such as flavonoids) in an Iranian breast cancer population. However, limitations are inevitable. No database for the flavonoids content of food is available in Iran. Diet is subject to changes after oncological diagnosis and related treatments. Since our FFQ was completed before surgery, this likely excluded cancer-related lifestyle modifications. Oncogenic metabolism is another variable that is imperative to unfold its effects on flavonoid turnover [1, 71]. The metabolism of flavonols might be affected under certain pathologic subtypes of breast tumors. Findings showed that higher plasma concentrations of flavonol aglycon (quercetin and kaempferol) were higher in hormone receptor-independent tumors. It might suggest that hormone-related signaling might be associated with the metabolism of flavonol aglycon which needs to be further clarified in experimental studies. Therefore validation analysis for polyphenols is suggested to be undergone considering pathological sub-classification. Although the test's accuracy has the privilege of the performance of a FFQ to indicate the dietary stratifications (high vs. low intake), the accuracy of cutoff points is an issue that remains as a limitation. The strength of the present study was testing median-, SR-, and ROC-based classifications to overcome this problem. The advantages of questioning in this case series study were dependent on the use of utensil images and disposable dishes to enhance the accuracy of the collected data. The present FFQ underwent face and content validities in a previous study [46]; besides, folate and cobalamin have been recently evaluated for the accuracy of this FFQ [48] [45, 46]. As for flavonols, folate is a nutrient that is predominantly found in plant-based food items [33].

Conclusions

Findings on the accuracy of a diagnostic test could support FFQ as a reasonable nutritional tool to estimate the intake status of quercetin and kaempferol among Iranian breast cancer patients. Despite that, the combined use of flavonol dietary data (FFQ) and biomarkers would allow nutritional epidemiologists to improve the estimators of the relationships between flavonols and health and guding clinicians in dietary advise.

Acknowledgments

We would like to give special thanks to Mrs. Parisa Varshosaz for helping to collect a part of the dietary data.

Funding

This is a part of project registered at Tabriz University of Medical Sciences, Tabriz, Iran (registration number D-97-5-701). This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. RZ-R was supported by the "Miguel Servet" program (CPII20/00009) from the Institute of Health Carlos III (Co-funded by the European Social Fund (ESF) - ESF investing in your future).

Declaration of interests

None.

Author contributions: RZ-R, VM, and SP: Conceptualization; S.Sad, RZ-R, VM, and SP: Data curation; RZ-R, VM, SS, PB, and SP: Formal analysis; S.Sad, VM, PB, and SP: Investigation; SP: Methodology; VM and SP: Project administration; SP: Resources; S.Sad, PB, and SP: Software; SP: Supervision; RZ-R, SS, and SP: Validation; SP: Visualization; RZ-R, VM, and SP: Roles/Writing - original draft; All authors: Writing - review & editing. All authors reviewed, edited and approved the final manuscript.

References

[1] Hollman PC. Absorption, bioavailability, and metabolism of flavonoids. Pharm Biol 2004; 42: 74-83.

[2] Pirouzpanah S, Hanaee J, Razavieh S-V, Rashidi M-R. Inhibitory effects of flavonoids on aldehyde oxidase activity. J Enzyme Inhib Med Chem 2009; 24: 14-21. 10.1080/14756360701841301.

[3] Jafarpour-Sadegh F, Montazeri V, Adili A, Esfehani A, Rashidi M-R, Pirouzpanah S. Consumption of fresh yellow onion ameliorates hyperglycemia and insulin resistance in breast cancer patients during doxorubicin-based chemotherapy: a randomized controlled clinical trial. Integr Cancer Ther 2017; 16: 276-89. 10.1177/1534735416656915.

[4] Manach C, Williamson G, Morand C, Scalbert A, Rémésy C. Bioavailability and bioefficacy of polyphenols in humans. Am J Clin Nutr 2005; 81: 230S-42S.

[5] Nijveldt RJ, Van Nood E, Van Hoorn DE, Boelens PG, Van Norren K, Van Leeuwen PA. Flavonoids: a review of probable mechanisms of action and potential applications. Am J Clin Nutr 2001; 74: 418-25.

[6] Adebamowo CA, Cho E, Sampson L, Katan MB, Spiegelman D, Willett WC, et al. Dietary flavonols and flavonol-rich foods intake and the risk of breast cancer. Int J Cancer 2005; 114: 628-33.

[7] Lee J, Mitchell AE. Pharmacokinetics of quercetin absorption from apples and onions in healthy humans. J Agric Food Chem 2012; 60: 3874-81. 10.1021/jf3001857.

[8] Zamora-Ros R, Rabassa M, Llorach R, González CA, Andres-Lacueva C. Application of dietary phenolic biomarkers in epidemiology: past, present, and future. J Agric Food Chem 2012; 60: 6648-57. 10.1021/jf204742e.

[9] Cao J, Zhang Y, Chen W, Zhao X. The relationship between fasting plasma concentrations of selected flavonoids and their ordinary dietary intake. Br J Nutr 2010; 103: d249-55. 10.1017/S000711450999170X.

[10] Erlund I, Silaste M, Alfthan G, Rantala M, Kesäniemi Y, Aro A. Plasma concentrations of the flavonoids hesperetin, naringenin and quercetin in human subjects following their habitual diets, and diets high or low in fruit and vegetables. Eur J Clin Nutr 2002; 56: 891-8. 10.1038/sj.ejcn.1601409.

[11] Iwasaki M, Inoue M, Sasazuki S, Miura T, Sawada N, Yamaji T, et al. Plasma tea polyphenol levels and subsequent risk of breast cancer among Japanese women: a nested case–control study. Breast Cancer Res Treat 2010; 124: 827-34. 10.1007/s10549-010-0916-x.

[12] Lan K, Jiang X, He J. Quantitative determination of isorhamnetin, quercetin and kaempferol in rat plasma by liquid chromatography with electrospray ionization tandem mass spectrometry and its application to the pharmacokinetic study of isorhamnetin. Rapid Commun Mass Spectrom 2007; 21: 112-20. 10.1002/rcm.2814.

[13] Linseisen J, Rohrmann S. Biomarkers of dietary intake of flavonoids and phenolic acids for studying diet–cancer relationship in humans. Eur J Nutr 2008; 47: 60-8. 10.1007/s00394-008-2007-x.

[14] Block G, Hartman AM, Dresser CM, Carroll MD, Gannon J, Gardner L. A data-based approach to diet questionnaire design and testing. Am J Epidemiol 1986; 124: 453-69.

[15] Dabeek WM, Marra MV. Dietary quercetin and kaempferol: bioavailability and potential cardiovascular-related bioactivity in humans. Nutrients 2019; 11: 2288. 10.3390/nu11102288.

[16] Radtke J, Linseisen J, Wolfram G. Fasting plasma concentrations of selected flavonoids as markers of their ordinary dietary intake. Eur J Nutr 2002; 41: 203-9. 10.1007/s00394-002-0377-z.

[17] Song WO, Chun OK. Tea is the major source of flavan-3-ol and flavonol in the U.S. diet. J Nutr 2008; 138: 1543s-7s. 10.1093/jn/138.8.1543S.

[18] Nielsen SE, Kall M, Justesen U, Schou A, Dragsted LO. Human absorption and excretion of flavonoids after broccoli consumption. Cancer Lett 1997; 114: 173-4. 10.1016/s0304-3835(97)04654-5.
[19] 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. 10.1097/mco.0000000000000424.
[20] 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. 10.3945/ajcn.113.077743.
[21] Zhang Y, Cao J, Chen W, Yang J, Hao D, Zhang Y, et al. Reproducibility and relative validity of a food frequency questionnaire to assess intake of dietary flavonol and flavone in Chinese university campus population. Nutr Res 2010; 30: 520-6.
[22] Willett WC, Lenart E. Reproducibility and validity of food-frequency questionnaires. In: Willett WC, editor. Nutritional epidemiology. 2 ed. New York: Oxford University Press; 1998. p. 101-47.
[23] Gleason PM, Harris J, Sheean PM, Boushey CJ, Bruemmer B. Publishing nutrition research: validity, reliability, and diagnostic test assessment in nutrition-related research. Am Diet Assoc 2010; 110: 409-19. 10.1016/j.jada.2009.11.022.
[24] Fraser GE, Jaceldo-Siegl K, Henning SM, Fan J, Knutsen SF, Haddad EH, et al. Biomarkers of dietary intake are correlated with corresponding measures from repeated dietary recalls and food-frequency questionnaires in the Adventist Health Study-2. J Nutr 2016; 146: 586-94.
[25] Garro-Aguilar Y, Cayssials V, Achaintre D, Boeing H, Mancini FR, Mahamat-Saleh Y, et al. Correlations between urinary concentrations and dietary intakes of flavonols in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Eur J Nutr 2020; 59: 1481-92. 10.1007/s00394-019-02005-5.
[26] Burkholder-Cooley NM, Rajaram SS, Haddad EH, Oda K, Fraser GE, Jaceldo-Siegl K. Validating polyphenol intake estimates from a food-frequency questionnaire by using repeated 24-h dietary recalls and a unique method-of-triads approach with 2 biomarkers. Am J Clin Nutr 2017; 105: 685-94. 10.3945/ajcn.116.137174.
[27] Shim J-S, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health 2014; 36: e2014009.
[28] Wakai K. A review of food frequency questionnaires developed and validated in Japan. J Epidemiol 2009; 19: 1-11. 10.2188/jea.je20081007.
[29] Babić D, Sindik J, Missoni S. Development and validation of a self-administered food frequency questionnaire to assess habitual dietary intake and quality of diet in healthy adults in the Republic of Croatia. Coll Antropol 2014; 38: 1017-26.
[30] Burkholder-Cooley NM, Rajaram SS, Haddad EH, Oda K, Fraser GE, Jaceldo-Siegl K. Validating polyphenol intake estimates from a food-frequency questionnaire by using repeated 24-h dietary recalls and a unique method-of-triads approach with 2 biomarkers–3. Am J Clin Nutr 2017; 105: 685-94. 10.3945/ajcn.116.137174.
[31] Willett WC. Correction for the effects of measurement error. In: Willett WC, editor. Nutritional epidemiology. 2 ed. New York: Oxford University Press; 1998. p. 302-20.
[32] Willett WC. Issues in analysis and presentation of dietary data. In: Willett WC, editor. Nutritional epidemiology. 2 ed. New York: Oxford University Press; 1998. p. 321-46.
[33] Owens JE, Holstege DM, Clifford AJ. Comparison of two dietary folate intake instruments and their validation by RBC folate. J Agric Food Chem 2007; 55: 3737-40. 10.1021/jf063649h.
[34] Parise CA, Caggiano V. Breast cancer survival defined by the ER/PR/HER2 subtypes and a surrogate classification according to tumor grade and immunohistochemical biomarkers. J Cancer Epidemiol 2014; 2014: 1-11.
[35] de Ferrars RM, Czank C, Zhang Q, Botting NP, Kroon PA, Cassidy A, et al. The pharmacokinetics of anthocyanins and their metabolites in humans. Br J Pharmacol 2014; 171: 3268-82. 10.1111/bph.12676.
[36] Pourzand A, Tajaddini A, Pirouzpanah S, Asghari-Jafarabadi M, Samadi N, Ostadrahimi A-R, et al. Associations between dietary allium vegetables and risk of breast cancer: A hospital-based matched case-control study. J Breast Cancer 2016; 19: 292-300.
[37] Pirouzpanah S, Taleban F-A, Mehdipour P, Atri M, Foroutan-Ghaznavi M. Plasma total homocysteine level in association with folate, pyridoxine, and cobalamin status among Iranian primary breast cancer patients. Nutr Cancer 2014; 66: 1097-108. 10.1080/01635581.2014.948213.
[38] Chen AY, Chen YC. A review of the dietary flavonoid, kaempferol on human health and cancer chemoprevention. Food Chem 2013; 138: 2099-107. 10.1016/j.foodchem.2012.11.139.
[39] Williamson G, Kay CD, Crozier A. The bioavailability, transport, and bioactivity of dietary flavonoids: A review from a historical perspective. Compr Rev Food Sci Food Saf 2018; 17: 1054-112. 10.1111/1541-4337.12351.

[40] Wu Q, Odwin-Dacosta S, Cao S, Yager JD, Tang WY. Estrogen down regulates COMT transcription via promoter DNA methylation in human breast cancer cells. Toxicol Appl Pharmacol 2019; 367: 12-22. 10.1016/j.taap.2019.01.016.
[41] Chen WY, Colditz GA. Risk factors and hormone-receptor status: epidemiology, risk-prediction models and treatment implications for breast cancer. Nat Clin Pract Oncol 2007; 4: 415-23. 10.1038/ncponc0851.
[42] Lachat C, Hawwash D, Ocké MC, Berg C, Forsum E, Hörnell A, et al. Strengthening the reporting of observational studies in epidemiology–nutritional epidemiology (STROBE-nut): An extension of the STROBE statement. PLoS Med 2016; 13: e1002036. 10.1371/journal.pmed.1002036.
[43] Goldberg GR, Black AE, Jebb SA, Cole TJ, Murgatroyd PR, Coward WA, et al. Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. Eur J Clin Nutr 1991; 45: 569-81.
[44] Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. Hum Nutr Clin Nutr 1985; 39: 5-41.
[45] Pirouzpanah S, Taleban F-A, Mehdipour P, Atri M, Hooshyareh-Rad A, Sabour S. The biomarker-based validity of a food frequency questionnaire to assess the intake status of folate, pyridoxine and cobalamin among Iranian primary breast cancer patients. Eur J Clin Nutr 2014; 68: 316-23. 10.1038/ejcn.2013.209.
[46] Pirouzpanah S, Taleban F-A, Sabour S, Mehdipour P, Atri M, Farrin N, et al. Validation of food frequency questionnaire to assess folate intake status in breast cancer patients. Razi J Med Sci 2012; 18: 31-41.
[47] Pirouzpanah S, Taleban FA, Abadi AR, Atri M, Mehdipour P. The association of plasma folate, vitamin B12 and homocysteine levels on hypermethylation status of rarβ2 gene in primary breast carcinoma. Iran J Epidemiol; 5: 19-27.
[48] Pirouzpanah S, Varshosaz P, Fakhrjou A, Montazeri V. The contribution of dietary and plasma folate and cobalamin to levels of angiopoietin-1, angiopoietin-2 and Tie-2 receptors depend on vascular endothelial growth factor status of primary breast cancer patients. Sci Rep 2019; 9: 14851. 10.1038/s41598-019-51050-x.
[49] Bhagwat S, Haytowitz DB, Wasswa-Kintu S. USDA Database for the Flavonoid Content of Selected Foods Release 3.2 [Internet]. Nutrient Data Laboratory, Beltsville Human Nutrition Research Center, ARS, USDA. 2015. Available from: http://www.ars.usda.gov/nutrientdata/flav. 10.15482/USDA.ADC/1178142
[50] Ader P, Wessmann A, Wolffram S. Bioavailability and metabolism of the flavonol quercetin in the pig. Free Radic Biol Med 2000; 28: 1056-67.
[51] Vashist SK, Luong JHT. Bioanalytical requirements and regulatory guidelines for immunoassays. In: Vashist SK, Luong JHT, editors. Handbook of immunoassay technologies: approaches, performances, and applications. Cambridge, Massachusetts: Academic Press; 2018. p. 81-95.
[52] Shrivastava A, Gupta VB. Methods for the determination of limit of detection and limit of quantitation of the analytical methods. Chron Young Sci 2011; 2: 21-5.
[53] Grobbee DE, Hoes AW. Clinical Epidemiology: Principles, Methods, and Applications for Clinical Research. Burlington, MA, United State: Jones and Bartlett Publisher; 2015.
[54] Burak C, Brüll V, Langguth P, Zimmermann BF, Stoffel-Wagner B, Sausen U, et al. Higher plasma quercetin levels following oral administration of an onion skin extract compared with pure quercetin dihydrate in humans. Eur J Nutr 2017; 56: 343-53.
[55] Brantsæter AL, Haugen M, Rasmussen SE, Alexander J, Samuelsen SO, Meltzer HM. Urine flavonoids and plasma carotenoids in the validation of fruit, vegetable and tea intake during pregnancy in the Norwegian Mother and Child Cohort Study (MoBa). Public Health Nutr 2007; 10: 838-47. 10.1017/S1368980007339037.
[56] Stålhammar G. Validation of biomarkers and digital image analysis in breast pathology: Inst för onkologi-patologi/Dept of Oncology-Pathology; 2017.
[57] Noroozi M, Burns J, Crozier A, Kelly I, Lean M. Prediction of dietary flavonol consumption from fasting plasma concentration or urinary excretion. Eur J Clin Nutr 2000; 54: 143-9.
[58] Yuan C, Spiegelman D, Rimm EB, Rosner BA, Stampfer MJ, Barnett JB, et al. Relative Validity of Nutrient Intakes Assessed by Questionnaire, 24-Hour Recalls, and Diet Records as Compared With Urinary Recovery and Plasma Concentration Biomarkers: Findings for Women. Am J Epidemiol 2018; 187: 1051-63. 10.1093/aje/kwx328.
[59] Vernet C, Philippat C, Agier L, Calafat AM, Ye X, Lyon-Caen S, et al. An Empirical Validation of the Within-subject Biospecimens Pooling Approach to Minimize Exposure Misclassification in Biomarker-based Studies. Epidemiology 2019; 30: 756-67. 10.1097/ede.0000000000001056.
[60] Wang X, Zhao X, Gu L, Lv C, He B, Liu Z, et al. Simultaneous determination of five free and total flavonoids in rat plasma by ultra HPLC–MS/MS and its application to a comparative pharmacokinetic study in normal and hyperlipidemic rats. J Chromatogr B Analyt Technol Biomed Life Sci 2014; 953: 1-10. 10.1016/j.jchromb.2014.01.042.

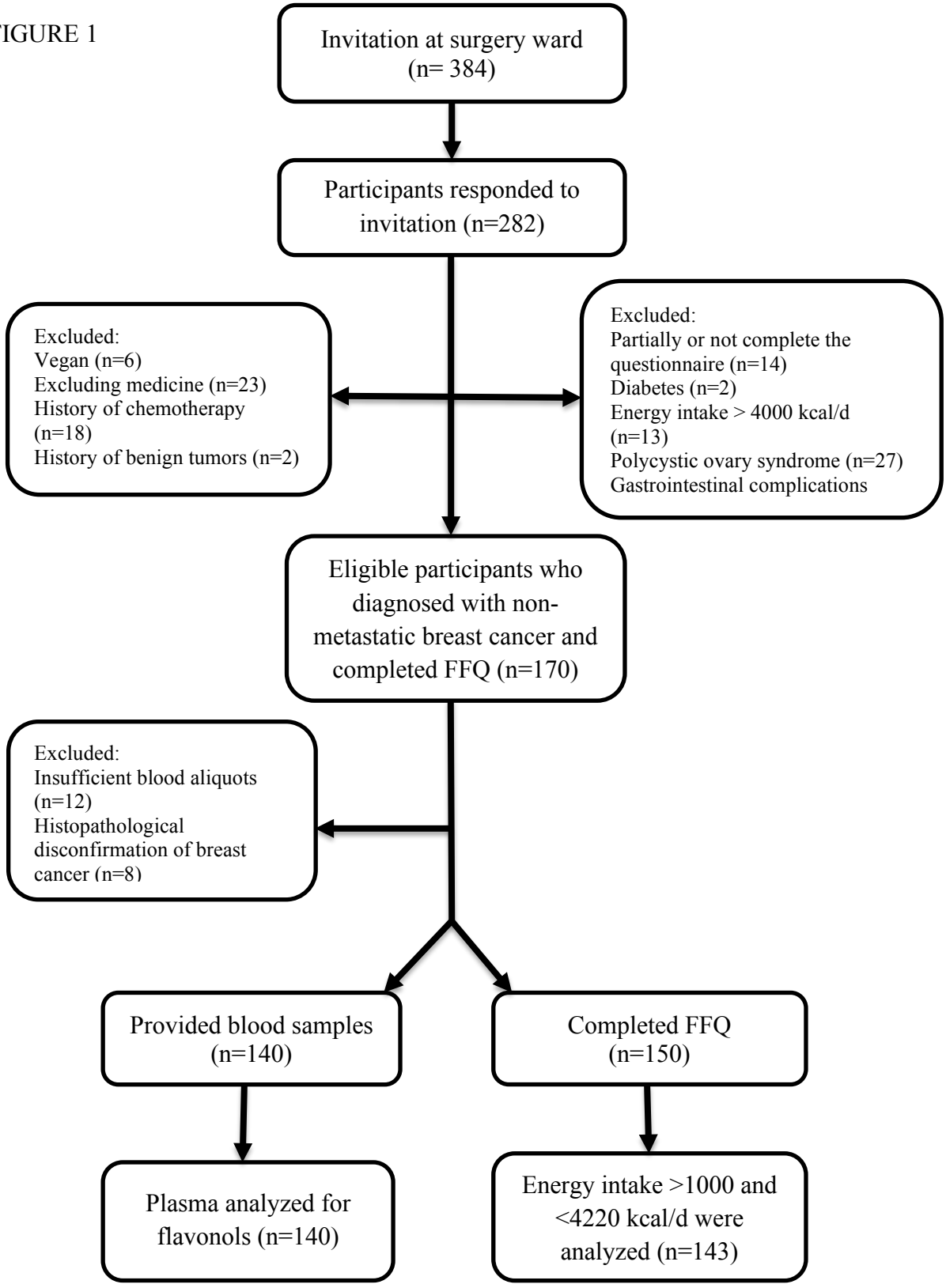
[61] Chan SG, Ho SC, Kreiger N, Darlington G, Adlaf EM, So KF, et al. Validation of a food frequency questionnaire for assessing dietary soy isoflavone intake among midlife Chinese women in Hong Kong. J Nutr 2008; 138: 567-73. 10.1093/jn/138.3.567.
[62] Fernandez AR, Omar SZ, Husain R. Development and validation of a food frequency questionnaire to estimate the intake of genistein in Malaysia. Int J Food Sci Nutr 2013; 64: 794-800. 10.3109/09637486.2013.798269.
[63] Nagata C, Nakamura K, Oba S, Hayashi M, Takeda N, Yasuda K. Association of intakes of fat, dietary fibre, soya isoflavones and alcohol with uterine fibroids in Japanese women. Br J Nutr 2009; 101: 1427-31. 10.1017/s0007114508083566.
[64] Verkasalo PK, Appleby PN, Allen NE, Davey G, Adlercreutz H, Key TJ. Soya intake and plasma concentrations of daidzein and genistein: validity of dietary assessment among eighty British women (Oxford arm of the European Prospective Investigation into Cancer and Nutrition). Br J Nutr 2001; 86: 415-21. 10.1079/bjn2001424.
[65] Wilkinson A, Gee J, Dupont M, Needs P, Mellon F, Williamson G, et al. Hydrolysis by lactase phlorizin hydrolase is the first step in the uptake of daidzein glucosides by rat small intestine in vitro. Xenobiotica 2003; 33: 255-64. 10.1080/0049825021000058088.
[66] Zhang Y, Li Y, Cao C, Cao J, Chen W, Zhang Y, et al. Dietary flavonol and flavone intakes and their major food sources in Chinese adults. Nutr Cancer 2010; 62: 1120-7. 10.1080/01635581.2010.513800.
[67] Sun C, Wang H, Wang D, Chen Y, Zhao Y, Xia W. Using an FFQ to assess intakes of dietary flavonols and flavones among female adolescents in the Suihua area of northern China. Public Health Nutr 2015; 18: 632-9. 10.1017/S1368980014000780.
[68] Swets J. Measuring the accuracy of diagnostic systems. Science 1988; 240: 1285-93. 10.1126/science.3287615.
[69] Carlsen MH, Karlsen A, Lillegaard IT, Gran JM, Drevon CA, Blomhoff R, et al. Relative validity of fruit and vegetable intake estimated from an FFQ, using carotenoid and flavonoid biomarkers and the method of triads. Br J Nutr 2011; 105: 1530-8. 10.1017/S0007114510005246.
[70] Zamora-Ros R, Ferrari P, González CA, Tjønneland A, Olsen A, Bredsdorff L, et al. Dietary flavonoid and lignan intake and breast cancer risk according to menopause and hormone receptor status in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study. Breast Cancer Res Treat 2013; 139: 163-76. 10.1007/s10549-013-2483-4.
[71] Katyal P, Bhardwaj N, Khajuria R. Flavonoids and their therapeutic potential as anticancer agents; biosynthesis, metabolism and regulation. World J Pharm Pharm Sci 2014; 3: 2188-216.

Legend to figures:

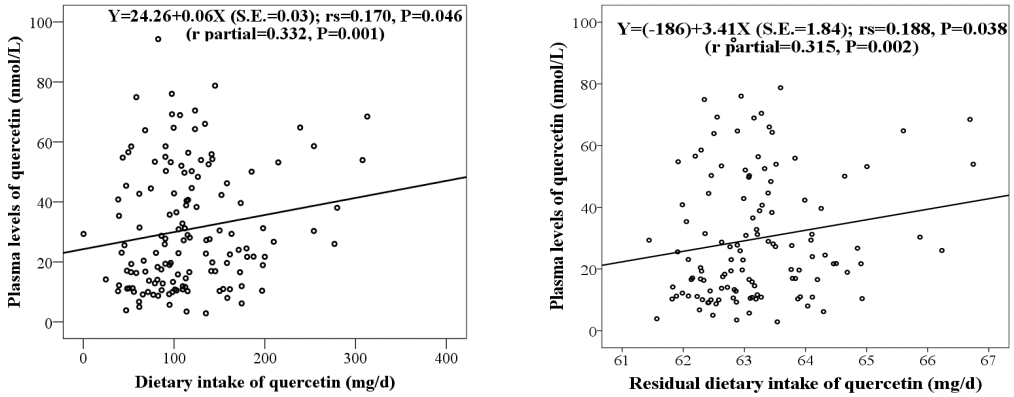
Fig. 1. Study design and flowchart diagram. Of the respondents (N=282) who were found to be eligible to enroll and participate, 140 women were included in the study.

Fig. 2. Scatter plots showing Spearman's rho correlation coefficients (rs) between daily dietary intake levels of flavonols (FFQ data, mg/d) and measured plasma level of certain flavonols (nmol/L) in breast cancer patients (N=140). Partial correlation analysis generated r adjusted for body mass index (kg/m^2), age at first menses (y), age at first pregnancy (y), and the number of breastfed child(ren). The equation of linear regression and the related standard error were also provided for each panel. There is also a correlation analysis between plasma concentrations of quercetin (independent variable) and isorhamnetin (dependent variable), which is shown in the last panel.

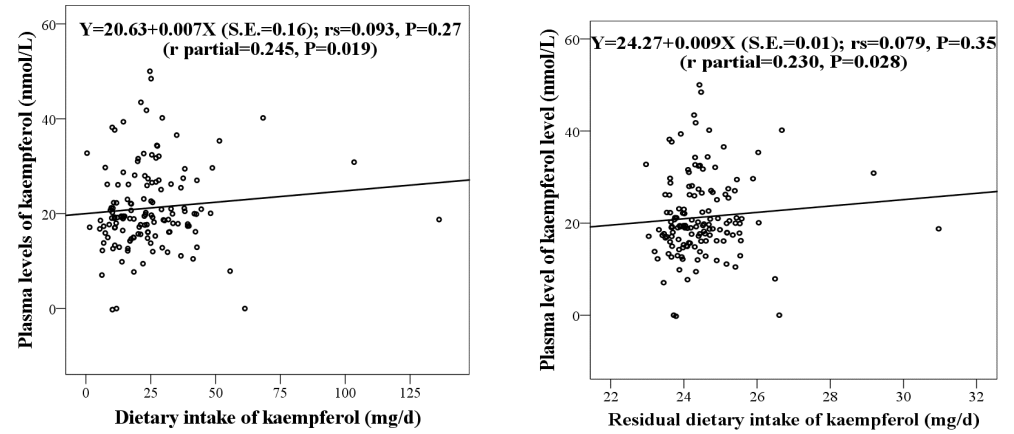
FIGURE 1


Invitation at surgery ward (n= 384)
Participants responded to invitation (n=282)
Excluded:
Vegan (n=6)
Excluding medicine (n=23)
History of chemotherapy (n=18)
History of benign tumors (n=2)
Excluded:
Partially or not complete the questionnaire (n=14)
Diabetes (n=2)
Energy intake > 4000 kcal/d (n=13)
Polycystic ovary syndrome (n=27)
Gastrointestinal complications
Eligible participants who diagnosed with non-metastatic breast cancer and completed FFQ (n=170)
Excluded:
Insufficient blood aliquots (n=12)
Histopathological disconfirmation of breast cancer (n=8)
Provided blood samples (n=140)
Completed FFQ (n=150)
Plasma analyzed for flavonols (n=140)
Energy intake >1000 and <4220 kcal/d were analyzed (n=143)

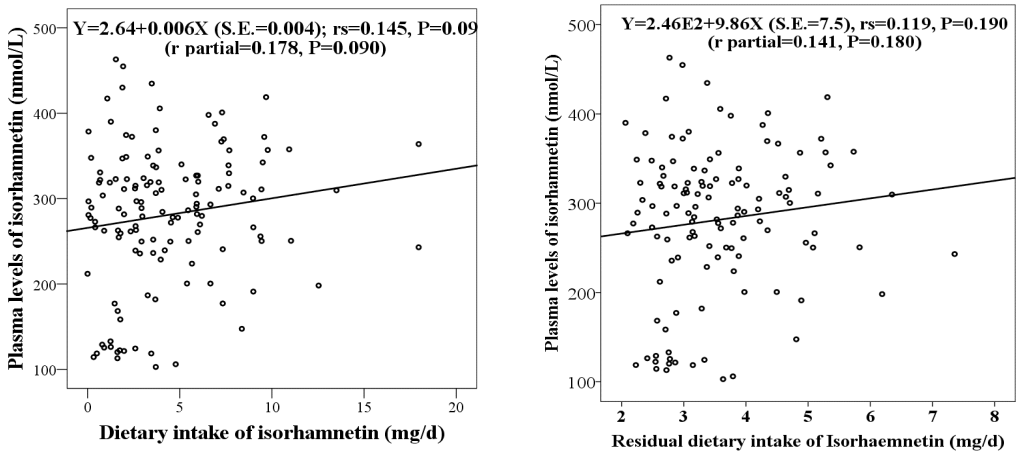
776 FIGURE 2



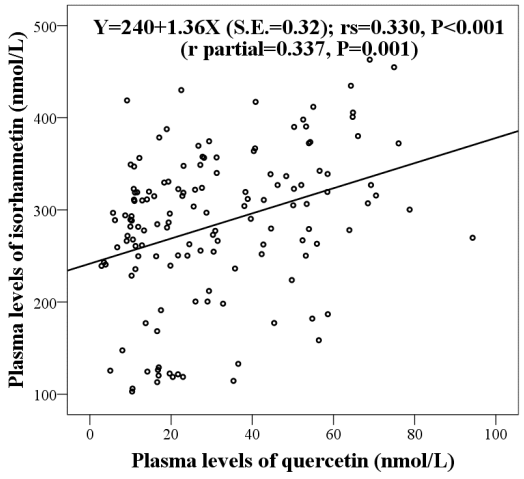
777



778



779



780

Table 1. General characteristics of women with breast cancer according to the median of plasma flavonol concentrations (N=140).

	Plasma concentrations of quercetin (nmol/L)					Plasma concentrations of kaempferol (nmol/L)					Plasma concentrations of isorhamnetin (nmol/L)				
Characteristic	**<85.9***		**≥85.9**			**<67.6***		**≥67.6**			**<928***		**≥928**		
Mean values	**Mean**	**SD**	**Mean**	**SD**	***P-value*** [a]	**Mean**	**SD**	**Mean**	**SD**	***P-value*** [a]	**Mean**	**SD**	**Mean**	**SD**	***P-value*** [a]
Age (years)	46.1	9.9	46.5	7.2	0.078	46.7	9.1	45.8	8.1	0.597	46.2	8.4	46.4	8.7	0.915
BMI (kg/m^2)	29.0	5.00	28.0	3.7	0.242	29.2	5.0	27.8	3.6	0.101	29.4	4.1	27.8	4.7	0.054
Age at first birth (years)	22.1	4.8	23.1	5.7	0.324	22.1	4.8	23.2	5.8	0.255	23.05	5.8	22.3	4.8	0.436
Total energy (kcal/day)	2744	945	2951	1331	0.329	2755	1079	2932	1254	0.404	2866	1123	2794	1208	0.735
Tumor size (cm)	3.2	1.5	2.9	1.5	0.361	3.0	1.5	3.1	1.5	0.629	3.2	1.5	2.9	1.4	0.220
Prevalence	**N**	**%**	**N**	**%**	***P-value*** [b]	**N**	**%**	**N**	**%**	***P-value*** [b]	**N**	**%**	**N**	**%**	***P-value*** [b]
Menopausal status															
Pre-menopause	48	70.6	58	81.7		49	71.0	57	81.4		58	82.9	49	70.0	
Post-menopause	20	29.4	13	18.3	0.124	20	29.0	13	18.6	0.149	12	17.1	21	30.0	0.073
Live birth (n)															
≤2	39	58.2	35	52.2		39	56.5	35	53.8		41	61.2	33	48.5	
>2	28	41.8	32	47.8	0.487	30	43.5	30	46.2	0.756	26	38.8	35	51.5	0.139
Breast-fed child (n)															
≤2	41	62.1	37	55.2		42	60.9	36	56.2		43	64.2	35	52.2	
>2	25	37.9	30	44.8	0.419	27	39.1	28	43.8	0.589	24	35.8	32	47.8	0.161
Smoking status															
Never smokers	19	90.5	25	100.0		18	90.0	26	100.0		22	91.7	22	100.0	
Ever smokers	2	9.5	0	0.0	0.115	2	10.0	0	0.0	0.099	2	8.3	0	0.0	0.166
Histological tumor grade															
I	12	19.7	7	11.3		12	19.3	7	11.5		11	18.0	8	2.9	
II	43	70.5	45	72.6		44	71.0	44	72.1		44	72.1	44	71.0	
III	6	9.8	10	16.1	0.308	6	9.7	10	16.4	0.315	6	9.9	10	16.1	0.481
Stage of breast cancer															
I	12	19.7	12	20.0		8	13.8	11	18.0		10	17.5	9	14.5	
II	43	70.5	11	18.3		16	27.6	15	24.6		13	22.8	19	30.6	
III	6	9.8	37	61.7	0.117	34	58.6	35	57.4	0.800	34	59.7	34	54.9	0.616
ALNM															
Positive	25	40.3	24	36.9		24	38.1	25	39.1		24	38.7	26	40.0	
Negative	37	59.7	41	63.1	0.694	39	61.9	39	60.9	0.911	38	61.3	39	60.0	0.882

SD, standard deviation; BMI, body mass index; ALNM, axillary lymph node metastasis.

[a] The *P-value* was obtained by an independent sample t-test.

[b] Chi-square test was performed.

Missing data: n=19 for menopausal status, n=88 for smoking status, n=19 for histopathological tumor grade, n=21 for the stage of breast cancer, n=12 for ALNM.

Table 2. Plasma concentrations and dietary intake of quercetin, kaempferol, and isorhamnetin among breast cancer patients (N=140).

	Total patients			Percentile		
Characteristics	**N**	**Mean**	**SD**	**25**	**50**	**75**
Plasma concentration (nmol/L)						
Quercetin	140	102.5	66.9	44.4	85.9	158.2
Kaempferol	140	74.3	30.7	56.4	67.6	91.6
Isorhamnetin	140	897	258	789	928	1059
Dietary data intake						
Total energy (kcal/day)	140	2624	809	1960	2543	3248
Quercetin (mg/day)	139	67.6	28.6	49.4	63.3	81.1
Kaempferol (mg/day)	139	24.4	17.0	13.6	22.5	30.3
Isorhamnetin (mg/day)	133	4.3	3.4	1.7	3.5	6.1

SD, standard deviation.

Table 3. Correlation coefficients between dietary intake of food groups (No.=10) and levels of intake or plasma concentrations of quercetin and kaempferol in the study population of women with breast cancer (n=140).

		Plasma concentrations of flavonols (nmol/L)						**Intake levels of Flavonols (mg/d)**					
	Food group	**Quercetin**		**Kaempferol**		**Isorhamnetin**		**Quercetin**		**Kaempferol**		**Isorhamnetin**	
No.	Group (g/day)	*r*	*β (S.E.)* [a]	*r*	*β (S.E.)* [a]	*r*	*β (S.E.)* [a]	*rs*	*β (S.E.)* [a]	*rs*	*β (S.E.)* [a]	*rs*	*β (S.E.)* [a]
1	Fruit	0.228**	0.11(0.10)	0.065	1.60(2.06)	0.059	87.0(147.5)	0.143	0.11(0.06)	0.219*	8.72(3.93)	0.250**	16.10(6.07)
	Fruit juice	0.068	3.97(4.98)	0.116	4.20(3.07)	*N.D.*	*N.D.*	0.149	22.77(14.00)	0.057	-1.92(6.00)	*N.D.* [b]	*N.D.* [b]
	Citrus	-0.040	-1.49(3.21)	0.047	2.69(4.93)	*N.D.*	*N.D.*	-0.066	-6.19(4.51)	0.134	10.00(9.54)	*N.D.* [b]	*N.D.* [b]
	Nectarine	-0.008	-0.02(0.19)	0.107	8.64(6.86)	*N.D.*	*N.D.*	0.179*	0.61(0.24)	0.062	14.32(13.34)	*N.D.* [b]	*N.D.* [b]
	Berries	0.198*	0.24 (0.14)	0.008	0.22(2.42)	0.025	190.4(647.5) [b]	0.292**	0.97(0.18)	0.093	7.96(4.66)	0.044	13.90(27.30) [b]
2	Vegetables	0.142	0.12(0.17)	0.241**	0.34(0.12)	0.143	3.57(2.12)	0.621**	1.60(0.16)	0.451**	1.04(0.22)	0.970**	1.01(0.21)
	Allium	0.135	0.15 (0.10)	0.085	0.64(0.64)	0.073	3.41(5.68)	0.636**	1.07(0.10)	0.405**	4.67(1.18)	0.410**	0.97(0.22)
	Cruciferus	0.187*	7.85(3.52)	0.191*	0.64(0.28)	*N.D.*	*N.D.*	0.218*	5.67(5.04)	0.274**	1.26(0.55)	*N.D.*	*N.D.*
	Apiaceous	0.077	1.86(2.07)	0.031	4.68(1.66)	0.122	32.4(22.56) [b]	0.216*	5.09(2.90)	0.259**	28.97(8.70)	0.930**	10.24(0.35) [b]
	Legumes	0.176**	5.37(2.57)	-0.045	-1.87(3.51)	*N.D.*	*N.D.*	0.112	21.50(7.14)	0.057	0.28(6.82)	*N.D.*	*N.D.*
3	Meats	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.*	*N.D.*	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]
4	Oil	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.*	*N.D.*	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]
5	Dried nuts	0.161	0.04(0.02)	-0.025	-2.01(7.00)	0.081	229.5(240.4) [b]	0.101	0.05(0.06)	0.120	11.36(13.53)	0.303**	14.30(10.08) [b]
6	Spices	0.122	4.76(3.30)	-0.017	-11.63(58.99)	0.057	875.9(1320.2) [b]	0.053	26.9(9.14)	0.081	13.3(114.4)	0.037	-4.94(55.50) [b]
7	Sweets	0.044	4.46(8.64)	-0.050	-10.04(17.15)	0.120	1812.9(1280.1) [b]	-0.023	-0.64(24.46)	-0.008	18.7(33.28)	0.032	10.98(54.34) [b]
8	Breads and cereals	-0.035	-0.08(0.20)	*N.D.* [b]	*N.D.* [b]	*N.D.*	*N.D.*	0.127	0.86(0.58)	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]
9	Milk and dairy products	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.*	*N.D.*	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]	*N.D.* [b]
10	Beverage	-0.042	-0.02(0.03)	0.004	0.01(0.05)	*N.D.*	*N.D.*	0.714**	0.63(0.08)	0.889**	1.01(0.03)	*N.D.* [b]	*N.D.* [b]

rs, Spearman's rho correlation coefficients; *r*, Pearson's correlation coefficients; S.E., standard error; *N.D.*, not determined.

[a] Simple linear regression analysis was used to estimate unstandardized β and standard error (S.E.).

[b] N.D. not detected. Data in the USDA database was zero or missing (49).

* Asterik indicates P<0.05 and ** for P<0.01.

Table 4. Area under ROC (receiver operating characteristic) curve and other diagnostic parameters were estimated to test the accuracy of dietary intake levels of quercetin and kaempferol (FFQ-base data, test measure) using the plasma concentrations of quercetin and kaempferol (as biochemical reference measure) of the study population of women with breast cancer (N=140).

Plasma concentartion *(nmol/L)*	**AUC**	**SE**	***P value***	**95% CI**	**Sensitivity (%)**	**Specificity (%)**
Total population						
Quercetin						
61.9 (ROC-based reference [a])	0.65	0.05	**0.004**	0.55-0.72	61.8	60.0
80.2 (SR [b])	0.61	0.05	**0.029**	0.51-0.70	79.4	39.4
85.9 (median [c])	0.60	0.05	**0.044**	0.50-0.69	60.9	60.0
Kaempferol						
60.1 (ROC-based reference [a])	0.60	0.05	0.078	0.49-0.70	59.4	60.5
57.8 (SR [b])	0.58	0.06	0.164	0.46-0.69	57.0	62.5
67.6 (median [c])	0.40	0.05	0.052	0.31-0.50	58.0	58.1
Pre-menopause						
Quercetin						
61.9 (ROC-based reference [a])	0.69	0.06	**0.004**	0.57-0.81	59.6	71.4
80.2 (SR [b])	0.65	0.06	**0.021**	0.53-0.77	78.0	54.3
85.9 (median [c])	0.63	0.06	**0.040**	0.51-0.75	57.4	65.8
Kaempferol						
60.1 (ROC-based reference [a])	0.64	0.07	**0.044**	0.49-0.74	58.7	72.7
57.8 (SR [b])	0.66	0.08	0.066	0.49-0.82	54.9	78.6
67.6 (median [c])	0.33	0.06	**0.007**	0.24-0.45	60.0	67.5
Post-menopause						
Quercetin						
61.9 (ROC-based reference [a])	0.62	0.1	0.211	0.43-0.81	68.4	53.9
80.2 (SR [b])	0.58	0.11	0.450	0.36-0.79	84.6	26.1
85.9 (median [c])	0.62	0.10	0.242	0.42-0.81	69.2	56.5
Kaempferol						
60.1 (ROC-based reference [a])	0.52	0.10	0.86	0.32-0.71	57.1	40.0
57.8 (SR [b])	0.53	0.10	0.79	0.33-0.72	59.1	42.8
67.6 (median [c])	0.64	0.10	0.15	0.46-0.83	42.9	40.9

ROC, receiver operating characteristic; N, number; AUC, area under the curve; SE, standard error; CI, confidence interval; SR, a standard reference.

[a] ROC-based reference model is describing the classification of plasma biomarkers based on cutoffs provided Youden's index which are supported by previous credential studies (quercetin ≥61.9 nmol/L, [10, 54]) and kaempferol ≥60.1 nmol/L [9]).

[b] The SR model is describing the classification of plasma biomarkers based on cutoffs provided by a previous credential study that is Cao et al. [9] research (quercetin ≥80.2 nmol/L and kaempferol ≥57.8 nmol/L).

[c] Median model is describing the classification of plasma biomarkers based on cutoffs provided by estimating median in the present study population (quercetin ≥85.9 nmol/L and kaempferol ≥67.6 nmol/L).

Supplementary Materials 1

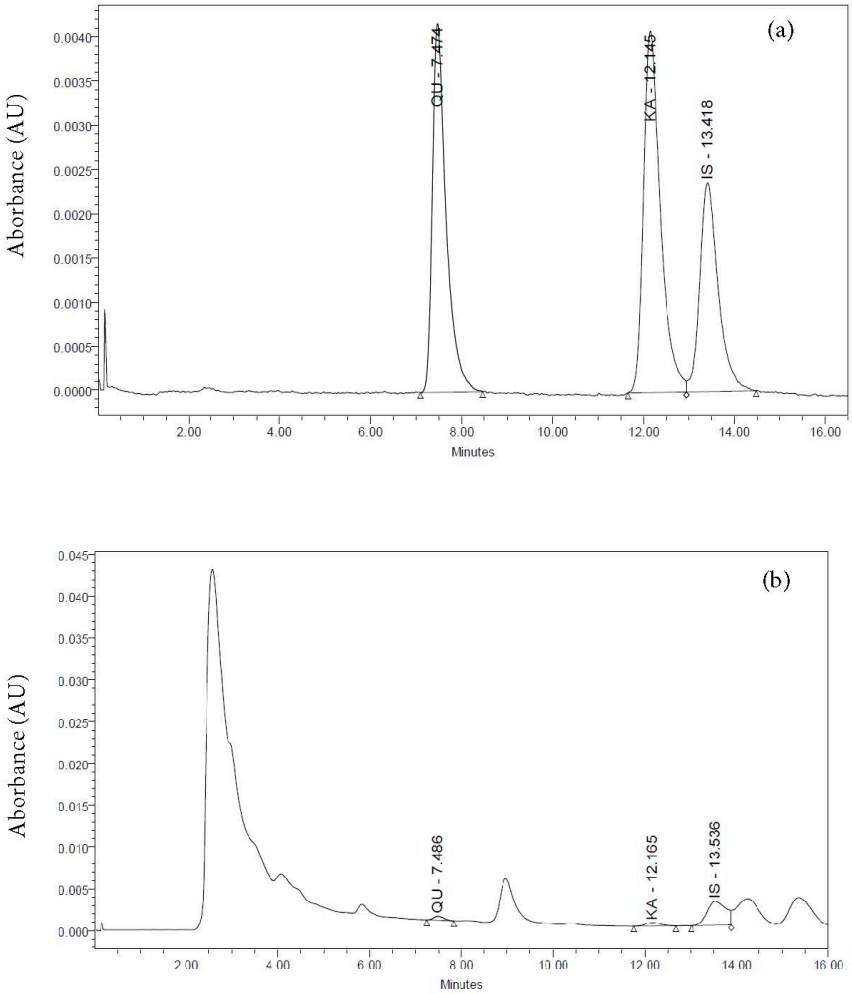
The content and face validities of present FFQ were detailed in the previous publication in Farsi [46]. Briefly, the Block's Health Habit and History Questionnaire (FFQ) [14] was translated into Persian by means of the standard method of "backward-forward" to develop basically a modified FFQ. At first, two independent bilingual experts translated into Persian. Backward translation into English was performed by a language institute where they collaborated with health-care centers for educational purposes. FFQ was harmonized with the forward translation. After reviewing the consistency, revising summarized dietary items by a group of breast cancer participants in a pilot set (n=25) to verify the importance, fluency, and understandability, the list of food items by two nutritional experts and English translators were reconciled, and FFQ-related questions contained 158 foods from locally available items were retained. Linguistic changes and required adjustments to improve the representativeness of food items were made based on the experts' comments and subsequently 145 food items remained. Face validity was undergone to verify reasonableness, appropriateness, attractiveness and the sequence of food items by a panel of experts (10 faculty members and health care professionals) at National Nutrition and Food Technology Research Institute, Shahid Beheshti University of Medical Sciences (Tehran, Iran). A three-point Likert scale entailing the necessity, helpful but not necessary, and not necessary were composed in another checklist for their review. The content validity rate (CVR) of questionnaire based on Lawshe table (above 0.62 as a criterion for 10 experts) to interpret the appropriateness and necessity of items [65]. The CVR assessment removed 9 items. The content validity (CVI) was obtained using Waltz and Bausell to examine the relevance, clarity, simplicity, and ambiguity of items of questionnaire based on four-point scale [66, 67]. All items were retained in questionnaire because the CVIs were above 0.75. A pilot sample of breast cancer patients (n=45) randomly selected from primary population in cancer clinic and asked for the clarity, simplicity of items listed in the 136-food item FFQ, and the five-point Likert scale to determine the importance of items from audiences' point of views [46]. This impact score was quantified, subsequently items gained >1.5 were retained in the questionnaire. The test-retest reliability was assessed by a secondary interview after two weeks for 20 women of study subjects individually who randomly assigned for this purpose. Pearson's correlation coefficient was obtained high

Legend to supplementary figure:

Supplementary Fig. 1. a) high-performance liquid chromatography (HPLC) chromatograms (370 nm) obtained for standard solutions prepared for quercetin, kaempferol, and isorhamnetin. **b)** A typical chromatogram is illustrating the retention time and area under the curve for flavonols extracted from plasma samples. **c)** The standard calibration curve was depicted using the area under the curve measured for certain standard concentrations of an analyte (flavonols). Linear equations ($y=ax^2+b$) obtained by regression analysis and correlation coefficient (R^2) for each standard calibration curve were also determined. Primarily, all data were measured in ng/ml and then converted to nmol/L in SPSS database.

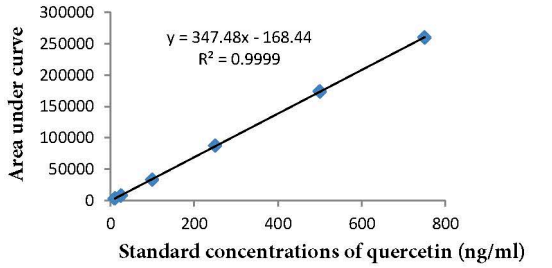
AU, absorbance units; QU, quercetin; IS, isorhamnetin; KA, kaempferol

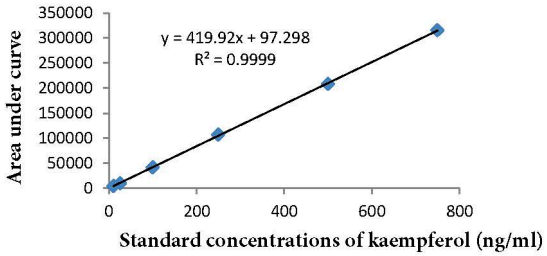
Supplementary Fig. 2. Bar diagrams comparing average plasma concentrations of flavonols (mean ± S.D.) between the dichotomous statuses of estrogen receptr (ER), progesterone receptor (PR), and human epidermal growth factor receptor (HER-2) in 140 breast cancer patients were tested using independent sample t-test. Luminal sub-classes were also applied to compare plasma concentrations of flavonols between groups using analysis of variance. Luminal A was considered as reference category when between pairwise comparisons were taken into account. Luminal A [ER+/PR+/HER-2(-)/histological grade 1, 2], luminal B [ER+/PR+/HER-2(+)/histological grade 3], triple-negative [ER-/PR-/HER-2(-)] and HER-2 positive [ER-/PR-/HER-2(+)] were classified based on recorded immunohistochemistry data of each participant [34, 56]. Asterisk sign (*) was used to express the significant level of a test ($P<0.05$).

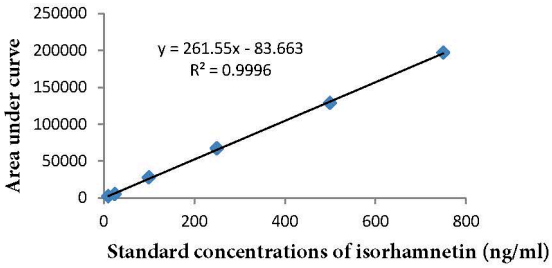


Supplement Fig. 1

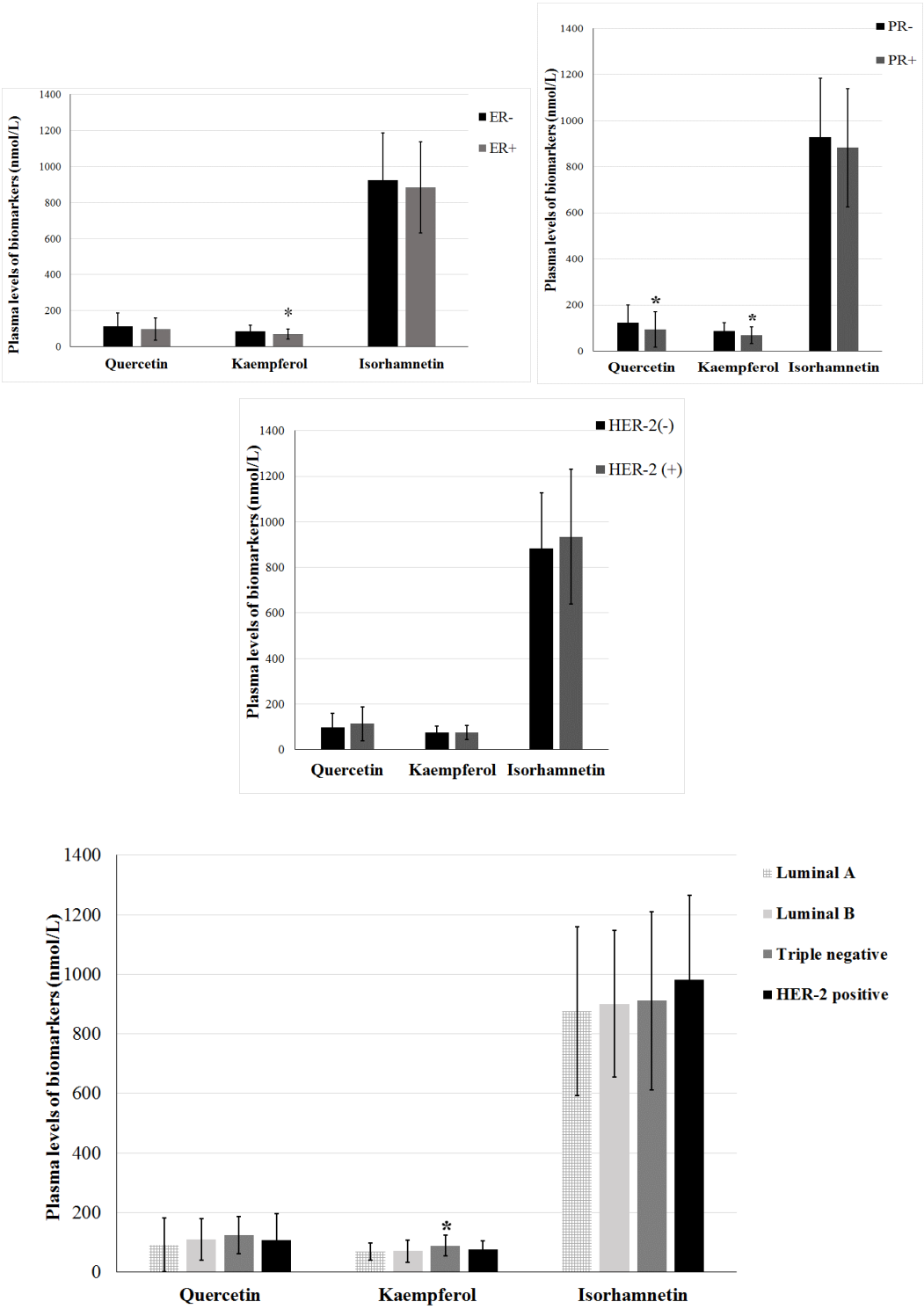
(c)







Supplement Fig. 1 (Cont'd)



Supplementary Fig. 2

Supplementary Table 1. STROBE-nut: An extension of the STROBE statement for nutritional epidemiology

Lachat C et al. (2016) STrengthening the Reporting of OBservational studies in Epidemiology – Nutritional Epidemiology (STROBE-nut): an extension of the STROBE statement. Plos Medicine 13(6) http://dx.doi.org/10.1371/journal.pmed.1002036 pdf or online version.

Item	Item nr	STROBE recommendations	Extension for Nutritional Epidemiology studies (STROBE-nut)	Reported on page #
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract.	**nut-1** State the dietary/nutritional assessment method(s) used in the title, abstract, or keywords.	**1, 2**
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found.		**2**
Introduction				
Background rationale	2	Explain the scientific background and rationale for the investigation being reported.		**3-4**
Objectives	3	State specific objectives, including any pre-specified hypotheses.		**3-4**
Methods				
Study design	4	Present key elements of study design early in the paper.		**4**
Settings	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection.	**nut-5** Describe any characteristics of the study settings that might affect the dietary intake or nutritional status of the participants, if applicable.	**4**

Item	Item nr	STROBE recommendations	Extension for Nutritional Epidemiology studies (STROBE-nut)	Reported on page #
Participants	6	a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up. Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls. Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants. (b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed. Case-control study—For matched studies, give matching criteria and the number of controls per case.	**nut-6** Report particular dietary, physiological or nutritional characteristics that were considered when selecting the target population.	**4-6**
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	**nut-7.1** Clearly define foods, food groups, nutrients, or other food components. **nut-7.2** When using dietary patterns or indices, describe the methods to obtain them and their nutritional properties.	**5-7** **N.A.**
Data sources - measurements	8	For each variable of interest, give sources of data and details of methods of assessment (measurement).Describe comparability of assessment methods if there is more than one	**nut-8.1** Describe the dietary assessment method(s), e.g., portion size estimation, number of days and items recorded, how it was developed and administered, and how quality was assured. Report if and how	**5-7**

Item	Item nr	STROBE recommendations	Extension for Nutritional Epidemiology studies (STROBE-nut)	Reported on page #
		group.	supplement intake was assessed.	
			nut-8.2 Describe and justify food composition data used. Explain the procedure to match food composition with consumption data. Describe the use of conversion factors, if applicable.	5-6
			nut-8.3 Describe the nutrient requirements, recommendations, or dietary guidelines and the evaluation approach used to compare intake with the dietary reference values, if applicable.	7
			nut-8.4 When using nutritional biomarkers, additionally use the STROBE Extension for Molecular Epidemiology (STROBE-ME). Report the type of biomarkers used and their usefulness as dietary exposure markers.	3 and 7
			nut-8.5 Describe the assessment of nondietary data (e.g., nutritional status and influencing factors) and timing of the assessment of these variables in relation to dietary assessment.	
			nut-8.6 Report on the validity of the dietary or nutritional assessment methods and any internal or external validation used in the study, if applicable.	5-6

Item	Item nr	STROBE recommendations	Extension for Nutritional Epidemiology studies (STROBE-nut)	Reported on page #
				1-16 (more specific: 5-6)
Bias	9	Describe any efforts to address potential sources of bias.	**nut-9** Report how bias in dietary or nutritional assessment was addressed, e.g., misreporting, changes in habits as a result of being measured, or data imputation from other sources	7
Study Size	10	Explain how the study size was arrived at.		**4-5**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why.	**nut-11** Explain categorization of dietary/nutritional data (e.g., use of N-tiles and handling of nonconsumers) and the choice of reference category, if applicable.	7
Statistical Methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions. (c) Explain how missing data were addressed. (d) Cohort study—If applicable, explain how loss to follow-up was addressed. Case-control study—If applicable, explain how	**nut-12.1** Describe any statistical method used to combine dietary or nutritional data, if applicable. **nut-12.2** Describe and justify the method for energy adjustments, intake modeling, and use of weighting factors, if applicable. **nut-12.3** Report any adjustments for measurement error, i.e,. from a validity or calibration study.	7 7 **More specific:** **At Table footnotes.**

Item	Item nr	STROBE recommendations	Extension for Nutritional Epidemiology studies (STROBE-nut)	Reported on page #
		matching of cases and controls was addressed. Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy. (e) Describe any sensitivity analyses.		**7 and Supplementary Fig. 1** **6** **7**
Results				
Participants	13	(a) Report the numbers of individuals at each stage of the study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed. (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram.	**nut-13** Report the number of individuals excluded based on missing, incomplete or implausible dietary/nutritional data.	**4-5 detailed in Fig. 1**
Descriptive data	14	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with	**nut-14** Give the distribution of participant characteristics across the exposure variables if applicable. Specify if food consumption of total population or consumers only were used to obtain results.	**7-8**

Item	Item nr	STROBE recommendations	Extension for Nutritional Epidemiology studies (STROBE-nut)	Reported on page #
		missing data for each variable of interest (c) Cohort study—Summarize follow-up time (e.g., average and total amount)		**4-5**
Outcome data	15	Cohort study—Report numbers of outcome events or summary measures over time. Case-control study—Report numbers in each exposure category, or summary measures of exposure. Cross-sectional study—Report numbers of outcome events or summary measures.		**N.A.**
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included. (b) Report category boundaries when continuous variables were categorized. (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period.	**nut-16** Specify if nutrient intakes are reported with or without inclusion of dietary supplement intake, if applicable.	**8-9**
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions and sensitivity analyses.	**nut-17** Report any sensitivity analysis (e.g., exclusion of misreporters or outliers) and data imputation, if applicable.	**8-9**

Item	Item nr	STROBE recommendations	Extension for Nutritional Epidemiology studies (STROBE-nut)	Reported on page #
Discussion				
Key results	18	Summarize key results with reference to study objectives.		**9-11**
Limitation	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	**nut-19** Describe the main limitations of the data sources and assessment methods used and implications for the interpretation of the findings.	**10-11**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	**nut-20** Report the nutritional relevance of the findings, given the complexity of diet or nutrition as an exposure.	**9-11**
Generalizability	21	Discuss the generalizability (external validity) of the study results.		**9-11**
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based.		**12**
Ethics			**nut-22.1** Describe the procedure for consent and study approval from ethics committee(s).	**5**
Supplementary material			**nut-22.2** Provide data collection tools and data as online material or explain how they can be accessed.	**18 & Data availability statement (Supplement**

Item	Item nr	STROBE recommendations	Extension for Nutritional Epidemiology studies (STROBE-nut)	Reported on page #
				information)

Supplementary Table 2. Information (source data) included in the calculation of sample size using the mean comparison formula by considering the different assumptions made are presented. Other data kept constant for calculation is the level of significance at 95% (α=0.05), and β=0.2.

		Variability in the source data					Type of calculation
	Assumption	**Method**	Mean1	SD1	Mean2	SD2	Sample size*
1) Total flavonoids (mg/d)		FFQ					**Average sample size to study FLAVONOIDS=571**
Song et al. (2008) [68]	Tea consumers vs. Nonconsumers		32.6	1.5	697.9	28.1	$n = \frac{(1.96 \times 0.84)^2 \times (2.25 + 789.6)}{598^2} = 0.017$
Nimptsch et al. (2016) [69]	Case (colorectal cancer); Quintile 5 (Q5) vs. Q1 HPFS+NHS cohorts	116-131 item-FFQ	116	36	769	355	$n = \frac{(1.96 \times 0.84)^2 \times (1296 + 126025)}{587.7^2} = 2.89$
Gates et al. (2009) [70]	Case (ovarian cancer)-control	FFQ	6.0	1.65	27.5	18.675	$n = \frac{(1.96 \times 0.84)^2 \times (2.72 + 348.75)}{19.35^2} = 7.36$
Real et al. (2018) [71]	Case (prostate cancer)-control	110 item-FFQ	286.0	207.4	268.7	166.9	$n = \frac{(1.96 \times 0.84)^2 \times (43006 + 27845)}{15.6^2} = 2275$
2) Dietary Flavonols (mg/d)							**Average sample size to study FLAVONOLS~67.2**
Real et al. (2018) [71]	Case (prostate cancer)-control	110 item-FFQ	63.36	46.85	37.14	29.23	$n = \frac{(1.96 \times 0.84)^2 \times (2195 + 854)}{23.6^2} = 49.3$
Zhang et al. (2010) [72]	Healthy subjects, between age groups	126 item-FFQ	13.05	5.07	15.39	7.16	$n = \frac{(1.96 \times 0.84)^2 \times (51.3 + 25.7)}{5.47} = 110.2$
Culter et al. (2007) [73]	Different cancers-Quintile 5 (Q5) vs. Q1 Iowa Women's' Health	127 item-FFQ	4.1	1.3**	21	27.7	$n = \frac{(1.96 \times 0.84)^2 \times (1.69 + 767.29)}{15.21^2} = 26.06$
Nimptsch et al. (2016) [69]	Case (colorectal cancer); Quintile 5 (Q5) vs. Q1 **HPFS** cohorts	116-131 item-FFQ	10.5	6.5	32	14.3	$n = \frac{(1.96 \times 0.84)^2 \times (42.25 + 204.49)}{19.35^2} = 5.16$
Nimptsch et al. (2016) [69]	Case(colorectal cancer); Quintile 5 (Q5) vs. Q1 **NHS** cohorts	116-131 item-FFQ	9.1	5.8	31.9	13.6	$n = \frac{(1.96 \times 0.84)^2 \times (33.64 + 184.96)}{20.52^2} = 4.07$
Somerset et al. (2014) [74]	Healthy subjects (difference in flavonol intakes)	62-item FFQ	131.6	260.5			$n = \frac{(1.96 \times 0.84)^2 \times (2 \times 67860.25)}{17318.56} = 30.72$
Cassidy et al. (2014) [75]	Ovarian cancer (NHS II: Q5[high] vs. Q1[low])	FFQ	9.7	5.9	30.6	14	$n = \frac{(1.96 \times 0.84)^2 \times (34.8 \times 196)}{18.8^2} = 5.11$
Cassidy et al. (2014) [75]	Ovarian cancer (NHS: Q5 vs. Q1)	FFQ	10.2	6.2	29.6	13.13	$n = \frac{(1.96 \times 0.84)^2 \times (38.4 + 172.4)}{17.5^2} = 5.4$

Feng et al. (2019) [76]	Case (BrCa)-control	81 item-FFQ	39.64	21.95	35.83	22.36	$n = \frac{(1.96 \times 0.84)^2 \times (481 + 499)}{4.2^2} = 436$
Song et al. (2008) [68]	Tea consumers vs. Nonconsumers	NHANES dietary recalls	7.3	0.2	31	0.9	$n = \frac{(1.96 \times 0.84)^2 \times (0.04 + 0.81)}{21.3^2} = 0.014$
Average number of participants needed to study FLAVONOLS							
3) Dietary Quercetin (mg/d)							
Grinder-Pedersen et al. (2003) [77]	Organic consumers (**High** Flavonoids microg/10MJ) vs. conventional diet (**Low** Flavonoids)	Human crossover intervention study	2632	774	4198	1370	$n = \frac{(1.96 \times 0.84)^2 \times (599076 + 1876900)}{1409.4^2} = 9.77$
Gates et al. (2009) [70]	Case (ovarian cancer)-control	FFQ-126 items	8.5	5.2	9.1	6.4	$n = \frac{(1.96 \times 0.84)^2 \times (27.04 + 40.96)}{0.54^2} = 1828$
4) Dietary Kaempferol (mg/d)							
Grinder-Pedersen et al. (2003) [77]	Organic consumers (**High** Flavonoids microg/10MJ) vs. conventional diet (**Low** Flavonoids)	Human crossover intervention study	333	328	608	352	$n = \frac{(1.96 \times 0.84)^2 \times (107584 + 123904)}{247.5^2} = 29.6$
5) Dietary Isorhamnetin (mg/d)							
Grinder-Pedersen et al. (2003) [77]	Organic consumers (**High** Flavonoids microg/10MJ vs. conventional diet (**Low** Flavonoids)	Human crossover intervention study	496	93	0	327	$n = \frac{(1.96 \times 0.84)^2 \times (8649 + 106929)}{-446.4^2} = 4.5$
Isoflavonoids							
Frankenfeld et al. [78]	Case (benign breast disease and breast cancer)–control study	FFQ	50	69.8	39	51.4	$n = \frac{(1.96 \times 0.84)^2 \times (4872 + 2642)}{10^2} = 75.14$

* Mean difference (d) was considered ≤10% of the actual mean difference.

** Standard deviation was computed based on the range provided by the authors.

Supplementary Table 3. Cross-tabulation of dietary quercetin and kaempferol (FFQ, test outcome) against the plasma concentrations of biomarkers (flavonol) (N=140).

	Plasma concentrations of *flavonol* (nmol/L) SR model [a]					Plasma concentrations of *flavonol* (nmol/L) Median model [b]					Plasma concentrations of *flavonol* (nmol/L) ROC-based reference model [c]			
Dietary status (SR cutoffs)	**Low**	**High**	**Total**	***P-value***	**Dietary status (median cutoffs)**	**Low**	**High**	**Total**	***P-value***	**Dietary status (ROC-based cutoff)**	**Low**	**High**	**Total**	***P-value***
Total population														
Quercetin (mg/d)														
<52.0	26 [d]	15	41	**0.015**	< 63.3	42	27	69	**0.014**	< 61.7	30	34	64	**0.013**
≥52.0	40	58	98		≥ 63.3	28	42	70		≥ 61.7	20	55	75	
Kaempferol (mg/d)														
<22.0	20	46	66	0.052	< 22.5	41	29	70	0.051	< 21.1	26	39	65	**0.030**
≥22.0	12	61	73		≥ 22.5	29	40	69		≥ 21.1	17	57	74	
Pre-menopause														
Quercetin (mg/d)														
<52.0	19	11	30	**0.002**	< 63.3	25	20	45	**0.033**	< 61.7	20	23	43	**0.007**
≥52.0	16	39	55		≥ 63.3	13	27	40		≥ 61.7	8	34	42	
Kaempferol (mg/d)														
<22.0	11	32	43	**0.022**	< 22.5	27	18	45	**0.010**	< 21.1	16	26	42	**0.011**
≥22.0	3	39	42		≥ 22.5	13	27	40		≥ 21.1	6	37	43	
Post-menopause														
Quercetin (mg/d)														
<52.0	6	2	8	0.682[e]	< 63.3	13	4	17	0.137 [e]	< 61.7	9	6	15	0.194
≥52.0	17	11	28		≥ 63.3	10	9	19		≥ 61.7	8	13	21	
Kaempferol (mg/d)														
<22.0	6	9	15	0.908	< 22.5	9	8	17	0.342	< 21.1	6	9	15	0.864
≥22.0	8	13	21		≥ 22.5	13	6	19		≥ 21.1	9	12	21	

N, number; SR, standard reference; ROC, receiver operating characteristic.

[a] The SR model is describing the classification of plasma biomarkers based on cutoffs provided by the previous credential study that is Cao et al. [9] research (quercetin ≥80.2 nmol/L and kaempferol ≥57.8 nmol/L).

[b] Median model is describing the classification of plasma biomarkers based on cutoffs provided by estimating the median in the present study population (quercetin ≥85.9 nmol/L and kaempferol ≥67.6 noml/L).

[c] ROC-based reference model is describing the classification of plasma biomarkers based on cutoffs provided Youden's index by previous credential studies (quercetin ≥ 61.9 nmol/L, [10, 54] and and kaempferol ≥ 60.1 nmol/L [9]).

[d] Number of study subjects. **Menopause data was unknown for a few cases (unspecified data).** Therefore, the sum of data across menopausal status was not the same as what appeared in the total population. Unspecified data (based on Fig. 1) of index test or reference were excluded from the analysis.

[e] Fisher's exact test was performed.

Supplementary Table 4. Correlation coefficients between dietary intake of food groups specified by the method of preparation and the levels of intake or plasma concentrations of quercetin and kaempferol (nmol/L) in the study population of women with breast cancer (n=140).

		Quercetin				Kaempferol			
		Plasma concentrations (nmol/L)		Intake levels (mg/d)		Plasma concentrations (nmol/L)		Intake levels (mg/d)	
Food group	**Type of Preparation**	**r**	**r partial** [a]	**rs**	**r partial** [a]	**r**	**r partial** [a]	**rs**	**r partial** [a]
Food (g/day) [b]									
Vegetables	*Raw*	0.114	0.126	0.261*	0.271*	0.241*	0.301**	0.434**	0.343**
	Cooked	0.120	0.217*	0.573**	0.692**	0.046	0.133	0.046	0.394**
	Fried	0.102	0.114	0.187	0.201*	-0.078	0.045	0.011	0.012
	Pickle	0.007	0.010	0.210*	0.109	0.067	0.042	0.141	0.126
Fruits	*Raw*	0.090	0.146	0.269**	0.363**	0.104	0.055	0.198*	0.105
	Cooked	0.074	0.102	0.110	0.098	0.020	0.071	0.012	0.008
Beans	*Cooked*	0.150	0.187	0.049	0.036	0.045	0.009	0.057	-0.008
Hot drinks		0.036	0.012	0.741	0.569**	0.004	0.053	0.889**	0.927**

rs, Spearman's rho correlation coefficients; *r*, Pearson's correlation coefficients; S.E., standard error; N.D., not determined.
[a] r partial was adjusted for age at diagnosis (y),
[b] Some food groups have quite trace amounts or lacking flavonol contents and therefore no result was determined (N.D.).
* Asterik indicates P<0.05 and ** for P<0.01.