

Jakob Linseisen  
Sabine Rohrmann

# Biomarkers of dietary intake of flavonoids and phenolic acids for studying diet–cancer relationship in humans

and cancer risk is evaluated. In addition, available analytical methods are discussed with respect to the requirements for their integration in epidemiological studies, putting specific emphasis on the epidemiological validation of such markers. *Methods* The scientific literature was screened for epidemiologic studies on the relationship of flavonoid and phenolic acid concentrations in human specimens (i.e. blood, urine) and cancer risk. In addition, original data on intra- and inter-subject variability of several flavonoids and phenolic acids are presented. *Results* Although several techniques are used in bioavailability or short-term intervention studies, their integration in epidemiological studies is very limited. An exception are phytoestrogens where validated immunoassays allow the rapid measurement of large sample

numbers with small sample volume. For several polyphenols, the data on the epidemiologic validity encourages for their use in epidemiological studies. *Conclusions* There are valid possibilities for additional biomarkers of flavonoid and phenolic acid intake that are best applied in prospective studies with more than one biological sample per subject. Currently, a combination of a single biomarker measurement with long-term dietary intake estimates will probably be the most valuable choice to decrease measurement error in exposure data.

J. Linseisen (✉) · S. Rohrmann  
Unit of Nutritional Epidemiology  
Division of Cancer Epidemiology  
German Cancer Research Center  
Im Neuenheimer Feld 280  
69120 Heidelberg, Germany  
Tel.: +49-6221/22363  
Fax: +49-6221/22203  
E-Mail: j.linseisen@dkfz.de

## Introduction

Many thousands of different flavonoids and phenolic acids are found in plants, with major dietary sources including fruit, vegetables, tea, chocolate, and soy. Flavonoids are classified in the main sub-groups of flavonols, flavones, flavanols, flavanones, isoflavones, lignans, anthocyanins and proanthocyanins; phenolic

acids are attributed to two main subgroups, the hydroxybenzoic and the hydroxycinnamic acids [21]. Some polyphenols are found in a rather wide variety of foods (such as kaempferol occurring in many vegetables), whereas others are limited to only a few kinds of food (such as apigenin found in parsley and celery).

Polyphenols have long been suggested to be beneficial for maintaining health. Meanwhile, experimental studies confirmed several biological effects.

Amongst their health-promoting properties are anti-oxidant, antiviral, anti-inflammatory, and anti-cancer activities. Such chemopreventive agents can be effective at different stages of the carcinogenic process, both by blocking initiation and by suppressing the later stages involving promotion, progression, angiogenesis, invasion and metastasis. Several reviews have summarized the potential chemopreventive mechanisms for a number of polyphenolics [12, 23, 35, 42].

Based on the available experimental evidence, the conduct of observational studies (case-control studies and cohort studies) in humans is the next logical and necessary step towards confirmation of biological effects in vivo. For the ultimate proof of effect, intervention studies are necessary. Observational studies trying to explore the association between dietary factors and disease risk often assess the habitual dietary intake by means of food frequency questionnaires (FFQ). These measurements, however, suffer from imprecision [19], even more so when it comes to dietary components which are provided by only few kinds of foods that may be consumed only occasionally or only as ingredients (e.g. herbs). The problem is potentiated if bioavailability is moderate to low, as shown for various secondary plant products, or if metabolic activity of the intestinal microflora is required to form bioavailable metabolites (e.g. plant lignans). As a further limitation, estimates of dietary intake are hampered by incomplete or missing data in food composition tables.

The use of biomarkers of intake of such compounds should overcome some of these methodological problems in nutritional epidemiology. Analytical data obtained from measurements in human specimens are more precise, at least for the time point when the biological material was collected, and the measurement error is independent of that contained in questionnaire-based dietary intake data [16, 19]. It is obvious that the validity and reproducibility of analytical assays for measurement of nutritional biomarker needs to be demonstrated in advance of their use. However, the concept of epidemiological validity includes also the investigation of the stability of the marker, e.g. during storage of the biological specimens, and the intra- and inter-subject variation [38].

### Background: overview on bioavailability and metabolism of polyphenols

Polyphenols markedly differ from one another in their bioavailability and intestinal metabolism. Current evidence from bioavailability studies suggests that the bioavailability varies from about 0.3 to 43% (based on urinary excretion) of the dose administered (Table 1), reaching plasma concentrations of 0.02–4.0  $\mu\text{mol/l}$  at an intake level of 50 mg aglycone equivalents [22]. Bioavailability is determined by different factors, including the sugar moiety of the compound and its further metabolism by the gut microflora [21].

**Table 1** Pharmacokinetic data from 97 studies concerning the bioavailability of polyphenols (according to Manach et al. [22])

	$T_{\text{max}}$ (h)		$C_{\text{max}}$ ( $\mu\text{mol/l}$ )		Urinary excretion <sup>a</sup> (% of intake)		Elimination half-life (h)	
	Mean (SEM)	Range	Mean (SEM)	Range	Mean (SEM)	Range	Mean (SEM)	Range
Daidzin	6.3 (0.6)	4.0–9.0	1.92 (0.25)	0.36–3.14	42.3 (3.0)	21.4–62.0	5.3 (0.8)	3.4–8.0
Daidzein	4.9 (1.0)	3.0–6.6	1.57 (0.52)	0.76–3.00	27.5		8.5 (0.8)	7.7–9.3
Genistin	6.5 (0.6)	4.4–9.3	1.84 (0.27)	0.46–4.04	15.6 (1.8)	6.8–29.7	7.8 (0.7)	5.7–10.1
Genistein	4.1 (0.6)	3.0–5.2	2.56 (1.00)	1.26–4.50	8.6		7.1 (0.3)	6.8–7.5
Glycitin	5.0		1.88 (0.38)	1.50–2.26	42.9 (12.0)	19.0–55.3	8.9	
Hesperidin	5.5 (0.1)	5.4–5.8	0.46 (0.21)	0.21–0.87	8.6 (4.0)	3–24.4	2.2	
Naringin	5.0 (0.1)	4.6–5.5	0.50 (0.33)	0.13–1.50	8.8 (3.17)	1.1–30.2	2.1 (0.4)	1.3–2.7
Quercetingluc.	1.1 (0.3)	0.5–2.9	1.46 (0.45)	0.51–3.80	2.5 (1.2)	0.31–6.4	17.9 (2.2)	10.9–28.0
Rutin	6.5 (0.7)	4.3–9.3	0.20 (0.06)	0.09–0.52	0.7 (0.3)	0.07–1.0	19.9 (8.1)	11.8–28.1
(Epi)catechin	1.8 (0.1)	0.5–2.5	0.40 (0.09)	0.09–1.10	18.5 (5.7)	2.1–55.0	2.5 (0.4)	1.1–4.1
EGC	1.4 (0.1)	0.5–2.0	1.10 (0.40)	0.30–2.70	11.1 (3.5)	4.2–15.6	2.3 (0.2)	1.7–2.8
EGCG	2.3 (0.2)	1.6–3.2	0.12 (0.03)	0.03–0.38	0.06 (0.03)	0.0–0.1	3.5 (0.3)	2.5–5.1
Gallic acid	1.6 (0.2)	1.3–1.5	4.00 (0.57)	2.57–4.70	37.7 (1.0)	36.4–39.6	1.3 (0.1)	1.1–1.5
Chlorogenic acid	1.0		0.26		0.3			
Caffeic acid	1.4 (0.6)	0.7–2.0	0.96 (0.26)	0.45–1.35	10.7			
Ferulic acid	2.0		0.03		27.6 (17.6)	3.1–61.7		
Anthocyanins	1.5 (0.4)	0.7–4.0	0.03 (0.02)	0.001–0.20	0.4 (0.3)	0.004–5.1		
Proanthocyanidins	2.0		0.02 (0.01)	0.008–0.03				

All the data were converted corresponding to a supply of 50 mg aglycone equivalent

<sup>a</sup>Usually represent 24-h urine samples.  $T_{\text{max}}$ , time to reach  $C_{\text{max}}$ . AUC, area under the plasma concentration–time curve  
EGC epigallocatechin, EGCG epigallocatechin gallate

Isoflavones and gallic acid are polyphenols that are absorbed to the highest extent, followed by the catechins, flavanones and quercetin glucosides, although the kinetics described differs considerably. The proanthocyanidins, the galloylated tea catechins and the anthocyanins are much less absorbed. So far, data on the bioavailability of phenolic acids are limited. For cinnamic acid, a Na<sup>+</sup>-dependent carrier-mediated transport process has been described [42]. In contrast to the initial assumption of passive diffusion of the aglycone as the major transport form, carrier-mediated transport processes may be common in the absorption of polyphenols [42]. The most prominent example is the active transport of quercetin glucoside.

Flavonoids undergo extensive first-pass phase II metabolism in the intestinal epithelial cells and the liver, being substrates for methylation, sulfation, and glucuronidation. For many of the polyphenols, plasma half-life time is short, and baseline levels are reached within 24 h [20]. It can be assumed that a steady-state level of plasma polyphenols is only achievable through regular consumption of foods containing them. The plasma concentrations as found in free-living subjects (fasting status) are even lower as compared to the above-mentioned range measured after dietary intervention (see, for example, values in Table 2). Due to the rapid metabolism and excretion, urinary samples, ideally from 24-h sampling, could be the preferred specimen (higher absolute amounts of polyphenols after analytical extraction and enrichment) as compared to plasma samples. In epidemiological studies that included sampling of biological material for nutritional biomarker measurement in their study design, however, usually one blood sample (serum, plasma, white blood cells, red blood cells) and/or one spot urine sample, all of limited volume, were obtained at recruitment.

### Laboratory techniques for the analysis of polyphenols in biological specimens

Two reviews provide an overview of the analytical methods used to determine the structure and content of the flavonoids and phenolic acids contained in foods and in food-based matrices [24, 32]. In principle, these methods can be applied to human specimens with small modifications. Thus, only a brief overview of techniques and specific issues for the analysis of polyphenolic compounds in blood and urine is given in the following section.

**Hydrolysis** In foods, flavonoids are usually glycosylated and phenolic acids are ester-bound. Hydrolysis (acidic or enzymatic) is frequently used to simplify the analytical procedure, and the respective aglycones and free acids are subsequently detected

and quantified. Since also in human specimens flavonoids are largely found in methylated, sulphated, and glucuronidated form, enzymatic hydrolysis by means of sulfatase and  $\beta$ -glucuronidase is used [1, 2, 8, 40], unless a study aims at determining the exact metabolite structures. Phenolic acids undergo further metabolism and degradation, although part of the native compounds are often excreted unmodified via urine.

**Clean-up procedures** For human specimens, such as plasma, serum, or urinary samples, solid phase extraction (SPE)-columns provide the most convenient solution for removing matrix compounds that would otherwise disturb the analysis [1, 2, 8]. However, some techniques, such as immunoassays, may require only minor sample preparation [40].

**Separation and detection systems** The two major separation techniques for the quantification of polyphenolics are HPLC and GC-MS, although the use of LC-MS/MS is becoming increasingly common [1, 8, 24, 32]. For flavonoids, HPLC is the method of choice; the coupled detection systems include diode array detectors, mass-selective detectors as well as electrochemical or fluorometric detectors. Phenolic acids are often quantified by means of GC after derivatisation. Identification of compounds by means of mass fragmentation is used as a gold standard. However, a single mass-selective detector often fails to fulfill the requirements for sensitivity. Thus, also HPLC-ESI-MS-MS systems and similarly coupled devices have been used for analysis of polyphenols [8, 15].

The availability of antibodies for isoflavones and lignans has allowed for the development of antibody-based assays with a high degree of sensitivity [40]. For the purpose of quantification, the fluorescence emitted is recorded by means of a plate reader, with the option of time-resolved measurement. To the best of our knowledge, the scientific literature contains no report on use of the metabonomics techniques to characterize clusters of polyphenolic compounds that can potentially be used as biomarkers.

### Analytical validation

Analytical validity focuses on the ability of a test to measure accurately and reliably the biomarker of interest. The components of analytical validity are sensitivity and specificity, and test reliability [38].

**Specificity** When using mass-selective detection methods, the characteristic mass fragments are used to confirm the identity of the compounds. This differs from other methods, such as HPLC-UV/VIS, -electron capture detection (ECD), or use of fluorescence detectors in which peak identity cannot be confirmed with certainty [24, 32]. Antibody-based assays are also

subject to failures in specificity due to cross-reactions with other matrix compounds. For all these methods, confirmation of their results by use of MS-based techniques is necessary. However, MS-based systems may have problems at concentrations close to the detection limit in which the characteristic mass fragments may be absent.

**Sensitivity** Sensitivity is a major issue with regard to analytical methods for determining polyphenols in biological specimens. Particularly in epidemiological studies, in which the available sample volumes are usually very small, the sensitivity of a method can be decisive for whether it can be used or not (along with other factors such as analytical time and costs). Detection limits very close to 1 nmol/l of polyphenols have been reported for techniques involving HPLC-ECD, HPLC-MS-MS, LC-MS/MS, and TR-FIA (immunoassay) [2, 8, 40]. The sensitivity reported for HPLC-MS, LC-MS, GC-MS and HPLC with use of fluorometric detection is slightly lower [1, 24, 32].

**Laboratory validity and reliability** For each well-developed method, satisfactory figures concerning the analytical precision and accuracy are available. It is often described as coefficient of variation (CV) between repeated measurements of the same samples. Deviation from the “true” concentration is indicated as percentage recovery of added standard substances to a sample (“spiked” sample). The best results are obtained with MS-based techniques (CV <5–7%, recovery 90–105%) [1, 8, 15]; the immunoassays are located at the lower end of the scale with coefficients of variation close to or above 10% and recovery rates frequently at <90% [40]. However, working at the detection limit of a method is always a challenge, and data on the validity of the methods are usually obtained clearly above the detection limit.

## Epidemiological validation

Epidemiologic validation aims at characterizing the variability of a certain biomarker within the population. The main components of biomarker variability relevant for the conduct and interpretation of epidemiological studies are the biological variability both within a subject (intra-subject variation) and between subjects (inter-subject variation), variability due to measurement error (see above), and variability due to random error [38, 39]. Repeat samples from the study subjects over a longer time period are necessary (weeks, up to years) to describe intra- and inter-subject variation. In addition, potentially relevant information on the study participants that may influence inter-subject variation, and information on conditions under which samples have been collected,

stored, and analyzed (e.g. fasting status, season of blood sampling, batch of the assay, differential handling of specimens from cases and control, knowledge about the case-control status, etc.) should be collected [33]. Such information can possibly be used for adjustment or stratification in the statistical analysis in order to minimize measurement error. When studying the relationship of a certain biomarker with a health outcome, intra-group variability (e.g. within cases or within controls) should be minimized to allow for the identification of inter-group differences (e.g. difference between cases and controls), if they exist [38]. For illustration purposes, an example based on own laboratory data is given below.

### Example: intra- and inter-subject variation in plasma concentrations of selected polyphenolic compounds

#### ■ Enterolactone

Plasma enterolactone concentrations were measured by means of a time-resolved fluorescence immunoassay. Concerning the analytical validity, intra- and inter-assay coefficients of variation ranged between 3.1–6.1 and 6.1–8.6%, respectively. Using two plasma samples that were analyzed in each batch for quality control, a CV <10% was obtained [29]. For the purpose of assessing intra- and inter-subject variability, plasma enterolactone concentrations were analyzed in samples obtained in a small intervention study administering flaxseed over seven consecutive days (for further details see [7]). Three fasting blood samples were taken 1 week before starting the intervention, on the day when the intervention started and 2 weeks after the intervention phase. The average inter-individual CV in fasting plasma enterolactone concentrations was 97%. The CVs describing the variation within subjects are listed in Table 2. With exception of subject B, CV was below 50%, in three subjects even below 10%. This data indicates that a kind of steady-state plasma concentration can be reached that describes a subjects’ exposure to enterolignans. However, it is well described that the use of antibiotics drastically diminishes enterolactone concentrations over a longer time period [17]; this additional information has to be obtained from the subjects in order to be considered in the statistical analysis.

#### ■ Other flavonoids and phenolic acids

The above-mentioned study also measured plasma concentrations of a series of flavonoids and phenolic



**Table 2** Intra- and inter-individual variation in plasma concentrations of selected flavonoids and phenolic acids

Compound	Subject														Intra- Subject	Inter- Subject
	A		B		C		D		E		F		G			
	Mean (nmol/l)	CV	Mean (nmol/l)	CV	Mean (nmol/l)	CV	Mean (nmol/l)	CV	Mean (nmol/l)	CV	Mean (nmol/l)	CV	Mean (nmol/l)	CV		
Enterolactone**	6.0	0.07	0.5	1.73	10.7	0.33	2.8	0.21	5.4	0.02	1.4	0.04	18.8	0.48	0.41	0.97
Kaempferol	17.8	0.13	11.2	0.18	11.3	0.12	49.8	0.09	31.5	0.11	25.7	0.27	22.0	0.36	0.16	0.58
Quercetin	52.0	0.25	55.1	0.12	73.3	0.11	177.6	0.13	83.5	0.17	60.9	0.21	89.1	0.13	0.14	0.52
Isorhamnetin	11.1	0.60	3.7	0.32	9.8	0.49	22.7	0.23	9.7	0.33	8.7	0.19	8.6	0.25	0.30	0.64
Luteolin	14.1	0.82	24.6	0.13	20.0	0.12	70.5	0.55	33.4	0.40	34.6	0.32	29.6	0.07	0.30	0.62
Hesperetin	0 <sup>a</sup>	–	0 <sup>a</sup>	–	4.7	1.03	0 <sup>a</sup>	–	7.2	1.39	0 <sup>a</sup>	–	0 <sup>a</sup>	–	1.21	2.28
Gallocatechin	16.9	0.22	8.0	0.86	25.1	0.45	43.5	1.10	23.8	0.22	8.2	1.10	30.8	0.28	0.53	0.78
Caffeic acid	90.2	0.39	74.1	0.17	153.0	0.30	330.7	0.59	143.0	0.72	189.1	0.66	91.2	0.12	0.37	0.57
<i>p</i> -Coumaric acid	0 <sup>a</sup>	–	0 <sup>a</sup>	–	1.5	1.73	0 <sup>a</sup>	–	7.2	1.73	0 <sup>a</sup>	–	15.5	1.73	1.30	2.65
Ferulic acid	93.0	0.22	74.3	0.36	245.8	0.46	610.4	0.32	301.2	0.35	45.3	0.38	85.7	0.39	0.31	0.97
Vanillic acid	0 <sup>a</sup>	–	0 <sup>a</sup>	–	14.1	1.73	52.4	0.89	202.1	1.73	0 <sup>a</sup>	–	0 <sup>a</sup>	–	1.45	2.38
Syringic acid	65.7	0.88	49.5	0.44	81.71	0.70	107.2	0.63	69.0	0.07	17.3	1.73	219.4	0.06	0.56	0.86
Salicylic acid	15.9	1.73	29.5	1.73	86.6	0.46	– <sup>b</sup>	–	124.2	0.40	0 <sup>a</sup>	–	188.6	1.73	1.06	1.63
Ellagic acid	0 <sup>a</sup>	–	43.1	0.80	0 <sup>a</sup>	–	272.8	0.50	0 <sup>a</sup>	–	33.4	0.87	455.7	0.33	0.62	1.65

After overnight fasting three plasma samples were taken within a 4-week-period. For each subject and compound, mean plasma concentration (nmol/l) and coefficient of variation (CV) is given

<sup>a</sup>Below detection limit

<sup>b</sup>Use of acetylsalicylic acid

\* Mean coefficient of variation as a measure for intra- and inter-individual variation

\*\* Enterolacton measurements were performed by means of time-resolved fluorescence immunoassay [23]; a HPLC–ECD technique was used for the analysis of all other compounds [16]

acids. A detailed description of the HPLC–ECD based method (including enzymatic hydrolysis) and the results of the analytic validity are given elsewhere [2]. Briefly, recovery of added standard compounds varied between 81 and 106%. Intra- and inter-assay coefficients of variation ranged between 1.0–6.5 and 1.5–9.6%, respectively.

When applying this method to three fasting blood samples per subjects, the inter-individual coefficients of variation (CV) ranged between 52% (quercetin) and 265% (*p*-coumaric acid). The CVs describing the within-subject variation are listed in Table 2. For each compound, mean intra-subject variation was distinctly below the CV for inter-subject variation. Intra-subject variation was lowest for the flavonols quercetin and kaempferol, followed by isorhamnetin, the flavone luteolin, and the hydroxycinnamic acids ferulic acid and caffeic acid. An intra-subject CV above 100% was observed for the flavanones (hesperetin), some hydroxybenzoic acids (vanillic and salicylic acid), and *p*-coumaric acid. Thus, several compounds are possible candidates for biomarker use. In the case of flavonoids and phenolic acids, there are also good examples of further factors that affect validity of the measurements and should be carefully considered in the study design and statistical evaluation. Season is expected to distinctly affect the intake of certain foods rich in specific phenolic compounds;

thus, the extent of variation of biomarker concentrations by season needs to be investigated, and season of blood (or urine) collection should be considered in the statistical analysis. In addition, many polyphenolic compounds are subject to oxidative damage during sample handling and maybe also sample storage. Addition of antioxidants is recommended to increase stability of polyphenols during sample preparation and extraction [2].

## Relationship of biomarkers to dietary intake

### ■ Integration of biomarker measurements in bioavailability studies and other short-term interventional studies

Several short-term intervention studies demonstrated the applicability of biomarkers of polyphenolic compounds and confirmed the direct relationship between dietary intake and biomarker concentrations. A recent report summarized all scientific studies ( $n = 97$ ) that have been conducted so far on the bioavailability of polyphenols [22]. Most of the studies concerned only one or few compounds within a given subclass of polyphenols. Kinetic data from this report are summarized in Table 1 (according to ref. [22]). Relatively low plasma concentrations were

obtained in each case, even after the administration of polyphenol preparations or polyphenol-rich food corresponding to 50 mg aglycone equivalents. The large differences in bioavailability between the various compounds and classes of polyphenols are striking.

Also, a review of short-term intervention studies ( $n = 93$ ) in which polyphenols were administered (either as isolated compounds or in the form of foods or food extracts) to human subjects was published recently [41]. The studies included in this review examined the biological effects of polyphenols. Measurement of biomarkers of dietary intake was used to describe the internal dose of the compounds under investigation and partly also to control for adherence to the study protocol.

### ■ Integration of biomarker measurements in cross-sectional studies

Various studies have investigated the suitability of fasting plasma or urinary concentrations of polyphenols (mainly flavonols, flavanones or isoflavones) as biomarkers of polyphenol intake [2, 4, 6, 26, 27, 30, 31, 37]. Usually, plasma samples were taken after overnight fast and urine was collected as spot urine or over 24 h. The results of these studies suggest that the biomarker concentrations reasonably reflected short-term intake of the polyphenols under investigation, although one study failed to support this conclusion [6]. A fairly high variation in plasma polyphenol concentrations of free-living subjects following their habitual diet was described [30]. Statistically significant correlations between estimates of the dietary intake before blood sampling and fasting plasma concentrations of polyphenols (quercetin, kaempferol, naringenin, hesperetin) were reported with correlation coefficients of 0.30–0.46 and 0.42–0.64, respectively, when considering the diet over 1 week or 1 day before blood sampling [30]. However, it has to be pointed out that the validity of such correlations may be limited by the lack of precision in the estimates of dietary polyphenol intake. Due to the short half-life time of most polyphenols, steady-state plasma concentrations can only be achieved if the compounds are consumed

regularly, a precondition most likely to be fulfilled by compounds such as kaempferol that are widely distributed in plant foods. Urinary excretion of polyphenols in subjects on habitual diets was also shown to be significantly correlated with estimates of short-term intake of fruits and vegetables, the correlation coefficients for selected flavonols and flavanones ranging from 0.28 to 0.38 [26]. The urinary polyphenol excretion rates show a high degree of variability just as described for the plasma concentrations. To give an example of polyphenol excretion in 24-h urine samples of subjects on a habitual diet, Nielsen and coworkers [26] reported average (SD) concentrations of quercetin, kaempferol, naringenin, phloretin, and total flavonoids as being 25 (23), 50 (32), 701 (659), 76 (110), 1638 (1316)  $\mu\text{g}/24$  h, respectively.

Plasma and urinary polyphenol concentrations are not expected to reflect long-term or habitual dietary intake, although this has not been investigated extensively. One study reported correlation coefficients between 0.24 and 0.74 for plasma isoflavone concentrations and the dietary intake estimated from FFQ data [37].

### Integration of biomarkers into studies of diet and cancer risk

In large-scale epidemiological (etiological) studies on disease-related effects of dietary polyphenolic compounds, little use has been made of biomarker measurements. Hertog and colleagues [10] were the first to analyze commonly consumed foods in terms of their flavonol and flavone content by means of HPLC; their work provided the basis for intake estimations of dietary flavonols and flavones. In the following years, several studies on associations with the risk of cardiovascular disease or cancer of different sites were conducted, some of which with promising results [11]. Except for studies on phytoestrogens (isoflavones and lignans), the available literature provides only few reports in which biomarker measurements of polyphenolic compounds were applied (Table 3). In all four identified studies, measurements

**Table 3** Epidemiologic studies of the risk of cancer using biomarkers of dietary polyphenol intake (excluding isoflavones and lignans)

Author	Type	N, cases/controls	Specimen; compounds	Cancer site	Result
Dai et al. [3]	CCS	250/250	Urine; citrus flavonoids	Breast	n.s.
Zheng et al. [45]	CCS	60/60	Urine; total phenols	Breast	n.s.
Sun et al. [34]	Cohort: nested CCS	232/772	Urine; tea polyphenols	Gastric and esophageal	Significant inverse association
Yuan et al. [43]	Cohort: nested CCS	162/806	Urine; tea polyphenols	Colorectal	Significant inverse association for colon; n.s. for rectum

CCS case-control study, n.s., no significant association

were conducted in spot urine samples, analyzing for flavanones (citrus fruit) [3], catechins (tea polyphenols) [34, 43] or a summary measure of phenolic compounds [45]. The investigations were performed in the Shanghai Breast Cancer Study (case-control study) [3, 45], and in the prospective Shanghai Cohort Study [34, 43].

Biomarkers of phytoestrogen intake, however, were estimated in a series of studies using both plasma and urinary samples. Especially, the association with the risk of breast and prostate cancer was investigated. Concerning breast cancer, the identified scientific studies are summarized in Table 4, including six case-control studies and an equal number of case-control studies nested in cohort studies. The major reasons for the frequent measurement of biomarkers of phytoestrogen intake in epidemiological studies was the availability of immunoassays that are appropriate for studies with large sample numbers while the required sample volumes are small. Sophisticated and time-consuming methods often do not allow for analyzing a sufficient number of samples that would be required for ensuring sufficient statistical power. However, phytoestrogen analysis was performed by means of an isotope dilution liquid chromatography/tandem mass-spectrometry method in the work of Grace et al. [9] and Verheus et al. [36] (Table 4). Due to the concomitant measurement of several phytoestrogens in the same sample and the high analytic validity, this sophisticated method became time- and cost-efficient as compared to other available methods.

In view of the usually rather low sample volumes available in epidemiologic studies, the analysis of polyphenols is in most cases restricted to one or few compounds. However, analytical procedures that permit a variety of polyphenols to be determined in a single run were published recently [1, 2, 8, 15].

Many of the studies listed in Tables 3 and 4 were designed as case-control studies. An effect of diagnosis and treatment of cancer on the biomarker levels seems possible and could lead to biased results. Therefore, prospective cohort studies with biological material obtained years before onset of disease are the studies of choice for the application of biomarker measurements. In almost all large cohort studies around the world, sampling of blood or urine was performed only once, in most instances at recruitment of the study participants. However, biomarker measurement in repeated samples over time would distinctly decrease measurement error in exposure data. This is especially true for dietary compounds with a possibly high variability in intake, such as several flavonoids and phenolic acids.

## Conclusion

Certainly, biomarkers are promising in providing a more accurate and objective measure of dietary intake of polyphenolic compounds than estimates based on current or habitual dietary intake. However, due to the short half-life time of these substances, fairly regular dietary intake is necessary to achieve kind of steady-state levels in biological specimens. This depends upon the dietary habits of the investigated population, e.g. this shall work for isoflavone intake in Asian populations and maybe lignan or flavonol (kaempferol) intake in Western populations. For the use of biomarkers in epidemiological studies it is essential that their analytical and epidemiological validity have been investigated in depth. It should be kept in mind that in most cohort studies biological specimens are collected once, mostly at recruitment. Thus, it also appears that the biomarker approach will not provide a solution to all problems encountered

**Table 4** Epidemiological studies of the risk of breast cancer risk using biomarkers of dietary isoflavone intake and/or mammalian lignans

Author	Type	N, cases/controls	Specimen; compound(s)	Result
Ingram et al. [14]	CCS	144/144	Urine; equol, enterolactone	Significant inverse association
Zheng et al. [45]	CCS	60/60	Urine; isoflavones	n.s.
Murkies et al. [25]	CCS	18/20	Urine; isoflavones	Significant inverse association
Pietinen et al. [28]	CCS	194/208	Serum; enterolactone	Significant inverse association
Dai et al. [3]	CCS	250/250	Urine; isoflavones, lignans	Significant inverse association
Piller et al. [29]	CCS	220/237	Plasma; enterolactone	Significant inverse association
den Tonkelaar et al. [5]	Cohort: nested CCS	88/268	Urine; genistein, enterolactone	n.s.
Hulten et al. [13]	Cohort: nested CCS	248/492	Plasma; enterolactone	Significant positive association
Grace et al. [9]	Cohort: nested CCS	97/187,114/219	Serum, urine; isoflavones, lignans	Significant positive association for isoflavones
Kilkkinen et al. [18]	Cohort: nested CCS	206/215	Serum; enterolactone	n.s.
Zeleniuch-Jacquotte et al. [44]	Cohort: nested CCS	417/417	Serum; enterolactone	n.s.
Verheus et al. [36]	Cohort: nested CCS	383/383	Plasma; isoflavones, lignans	Significant inverse association for genistein levels; n.s. for lignans

CCS case-control study, n.s. no significant association

with (long-term) dietary intake calculations. A combination of methods will probably be the most valuable choice.

■ **Acknowledgments** This work was supported by Environmental Cancer Risk, Nutrition and Individual Susceptibility (ECNIS), a network of excellence operating within the European Union 6th Framework Program, Priority 5: "Food Quality and Safety" (Contract No 513943)

**Conflict of Interest:** none.

## References

- Adlercreutz H, Kiuru P, Rasku S, Wähälä K, Fotsis T (2004) An isotope dilution gas chromatographic-mass spectrometric method for the simultaneous assay of estrogens and phytoestrogens in urine. *J Steroid Biochem Mol Biol* 92:399–411
- Bolarinwa A, Linseisen J (2005) Validated application of a new high-performance liquid chromatographic method for the determination of selected flavonoids and phenolic acids in human plasma using electrochemical detection. *J Chromatogr B Analyt Technol Biomed Life Sci* 823:143–151
- Dai Q, Franke AA, Jin F, Shu XO, Hebert JR, Custer LJ et al (2002) Urinary excretion of phytoestrogens and risk of breast cancer among Chinese women in Shanghai. *Cancer Epidemiol Biomarkers Prev* 11:815–821
- de Vries JH, Hollman PC, Meyboom S, Buysman MN, Zock PL, van Staveren WA, Katan MB (1998) Plasma concentrations and urinary excretion of the antioxidant flavonols quercetin and kaempferol as biomarkers for dietary intake. *Am J Clin Nutr* 68:60–65
- den Tonkelaar I, Keinan-Boker L, Veer PV, Arts CJ, Adlercreutz H, Thijssen JH et al (2001) Urinary phytoestrogens and postmenopausal breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 10:223–228
- Erlund I, Silaste ML, Alftan G, Rantala M, Kesaniemi YA, Aro A (2002) 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* 56:891–898
- Fuchs D, Piller R, Linseisen J, Daniel H, Wenzel U (2007) The human peripheral blood mononuclear cell proteome responds to a dietary flaxseed-intervention and proteins identified suggest a protective effect in atherosclerosis. *Proteomics* 7:3278–3288
- Grace PB, Mistry NS, Carter MH, Leatham AJ, Teale P (2007) High throughput quantification of phytoestrogens in human urine and serum using liquid chromatography/tandem mass spectrometry (LC-MS/MS). *J Chromatogr B Analyt Technol Biomed Life Sci* 853:138–146
- Grace PB, Taylor JI, Low YL, Luben RN, Mulligan AA, Botting NP et al (2004) Phytoestrogen concentrations in serum and spot urine as biomarkers for dietary phytoestrogen intake and their relation to breast cancer risk in European prospective investigation of cancer and nutrition-Norfolk. *Cancer Epidemiol Biomarkers Prev* 13:698–708
- Hertog MG, Hollman PC, Katan MB, Kromhout D (1993) Intake of potentially anticarcinogenic flavonoids and their determinants in adults in The Netherlands. *Nutr Cancer* 20:21–29
- Hollman PC, Hertog MG, Katan MB (1996) Role of dietary flavonoids in protection against cancer and coronary heart disease. *Biochem Soc Trans* 24:785–789
- Howells LM, Moiseeva EP, Neal CP, Foreman BE, Andreadi CK, Sun YY, Hudson EA, Manson MM (2007) Predicting the physiological relevance of in vitro cancer preventive activities of phytochemicals. *Acta Pharmacol Sin* 28:1274–1304
- Hulten K, Winkvist A, Lenner P, Johansson R, Adlercreutz H, Hallmans G (2002) An incident case-referent study on plasma enterolactone and breast cancer risk. *Eur J Nutr* 41:168–176
- Ingram D, Sanders K, Kolybaba M, Lopez D (1997) Case-control study of phyto-oestrogens and breast cancer. *Lancet* 350:990–994
- Ito H, Gonthiera MP, Manach C, Morand C, Mennen L, Remesy C et al (2004) High-throughput profiling of dietary polyphenols and their metabolites by HPLC-ESI-MS-MS in human urine. *Biofactors* 22:241–243
- Kaaks R, Riboli E, Sinha R (1997) Biochemical markers of dietary intake. In: Toniolo P et al. (ed) *Application of biomarkers in cancer epidemiology*. IARC Scientific Publications No. 142, IARC, Lyon
- Kilkkinen A, Pietinen P, Klaukka T, Virtamo J, Korhonen P, Adlercreutz H (2002) Use of oral antimicrobials decreases serum enterolactone concentration. *Am J Epidemiol* 155:472–477
- Kilkkinen A, Virtamo J, Vartiainen E, Sankila R, Virtanen MJ, Adlercreutz H, Pietinen P (2004) Serum enterolactone concentration is not associated with breast cancer risk in a nested case-control study. *Int J Cancer* 108:277–280
- Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, Bingham S, Schoeller DA, Schatzkin A, Carroll RJ (2003) Structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol* 158:14–21; discussion 22–26
- Manach C, Donovan JL (2004) Pharmacokinetics and metabolism of dietary flavonoids in humans. *Free Radic Res* 38:771–785
- Manach C, Scalbert A, Morand C, Remesy C, Jimenez L (2004) Polyphenols: food sources and bioavailability. *Am J Clin Nutr* 79:727–747
- Manach C, Williamson G, Morand C, Scalbert A, Remesy C (2005) Bioavailability and bioefficacy of polyphenols in humans. I. Review of 97 bioavailability studies. *Am J Clin Nutr* 81(suppl 1):230S–242S
- Manson MM, Foreman BE, Howells LM, Moiseeva EP (2007) Determining the efficacy of dietary phytochemicals in cancer prevention. *Biochem Soc Trans* 35:1358–63
- Merken HM, Beecher GR (2000) Measurement of food flavonoids by high-performance liquid chromatography: a review. *J Agric Food Chem* 48:577–599
- Murkies A, Dalais FS, Briganti EM, Burger HG, Healy DL, Wahlqvist ML et al (2000) Phytoestrogens and breast cancer in postmenopausal women: a case control study. *Menopause* 7:289–296
- Nielsen SE, Freese R, Kleemola P, Mutanen M (2002) Flavonoids in human urine as biomarkers for intake of fruits and vegetables. *Cancer Epidemiol Biomarkers Prev* 11:459–466



27. Noroozi M, Burns J, Crozier A, Kelly IE, Lean ME (2000) Prediction of dietary flavonol consumption from fasting plasma concentration or urinary excretion. *Eur J Clin Nutr* 54:143–149
28. Pietinen P, Stumpf K, Mannisto S, Kataja V, Uusitupa M, Adlercreutz H (2001) Serum enterolactone and risk of breast cancer: a case-control study in eastern Finland. *Cancer Epidemiol Biomarkers Prev* 10:339–344
29. Piller R, Chang-Claude J, Linseisen J (2006) Plasma enterolactone and genistein and the risk of premenopausal breast cancer. *Eur J Cancer Prev* 15:225–232
30. Radtke J, Linseisen J, Wolfram G (2002) Fasting plasma concentrations of selected flavonoids as markers of their ordinary dietary intake. *Eur J Nutr* 41:203–209
31. Ritchie MR, Morton MS, Deighton N, Blake A, Cummings JH (2004) Plasma and urinary phyto-oestrogens as biomarkers of intake: validation by duplicate diet analysis. *Br J Nutr* 91:447–457
32. Robbins RJ (2003) Phenolic acids in foods: an overview of analytical methodology. *J Agric Food Chem* 51:2866–2887
33. Rundle AG, Vineis P, Ahsan H (2005) Design options for molecular epidemiology research within cohort studies. *Cancer Epidemiol Biomarkers Prev* 14:1899–1907
34. Sun CL, Yuan JM, Lee MJ, Yang CS, Gao YT, Ross RK, et al. (2002) Urinary tea polyphenols in relation to gastric and esophageal cancers: a prospective study of men in Shanghai, China. *Carcinogenesis* 23:1497–1503
35. Thomasset SC, Berry DP, Garcea G, Marczylo T, Steward WP, Gescher AJ (2007) Dietary polyphenolic phytochemicals—promising cancer chemopreventive agents in humans? A review of their clinical properties. *Int J Cancer* 120:451–458
36. Verheus M, van Gils CH, Keinan-Boker L, Grace PB, Bingham SA, Peeters PH (2007) Plasma phytoestrogens and subsequent breast cancer risk. *J Clin Oncol* 25:648–655
37. Verkasalo PK, Appleby PN, Allen NE, Davey G, Adlercreutz H, Key TJ (2001) 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* 86:415–421
38. Vineis P, Gallo V (2007) The epidemiological theory: principles of biomarker validation. In: Vineis P, Gallo V (eds) *Epidemiological concepts of validation of biomarkers for the identification/quantification of environmental carcinogenic exposures*. ECNIS publications 3, NOFER Institute, Lodz, pp 9–18
39. Vineis P, Perera F (2007) Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old. *Cancer Epidemiol Biomarkers Prev* 16:1954–1965
40. Wang GJ, Lapcik O, Hampl R, Uehara M, Al-Maharik N, Stumpf K et al (2000) Time-resolved fluoroimmunoassay of plasma daidzein and genistein. *Steroids* 65:339–348
41. Williamson G, Manach C (2005) Bioavailability and bioefficacy of polyphenols in humans. II. Review of 93 intervention studies. *Am J Clin Nutr* 81(suppl 1):243S–255S
42. Yang CS, Landau JM, Huang MT, Newmark HL (2001) Inhibition of carcinogenesis by dietary polyphenolic compounds. *Annu Rev Nutr* 21:381–406
43. Yuan JM, Gao YT, Yang CS, Yu MC (2007) Urinary biomarkers of tea polyphenols and risk of colorectal cancer in the Shanghai Cohort Study. *Int J Cancer* 120:1344–1350
44. Zeleniuch-Jacquotte A, Adlercreutz H, Shore RE, Koenig KL, Kato I, Arslan AA, Toniolo P (2004) Circulating enterolactone and risk of breast cancer: a prospective study in New York. *Br J Cancer* 91:99–105
45. Zheng W, Dai Q, Custer LJ, Shu XO, Wen WQ, Jin F et al (1999) Urinary excretion of isoflavonoids and the risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 8:35–40