

## Frequency of soy food consumption and serum isoflavone concentrations among Chinese women in Shanghai

Cara L Frankenfeld<sup>1,2</sup>, Johanna W Lampe<sup>1,2,\*</sup>, Jackilen Shannon<sup>3</sup>, Dao L Gao<sup>4</sup>, Roberta M Ray<sup>1</sup>, Joann Prunty<sup>1</sup>, Thomas F Kalhorn<sup>5</sup>, Kristiina Wähälä<sup>6</sup>, Ruth E Patterson<sup>1,2</sup> and David B Thomas<sup>1,2</sup>

<sup>1</sup>Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North M4-B402, Seattle, WA 98109, USA: <sup>2</sup>Department of Epidemiology, University of Washington, Seattle, WA, USA: <sup>3</sup>Portland Veterans Administration Medical Center, Portland, OR, USA: <sup>4</sup>Department of Epidemiology, Zhong Shan Hospital Cancer Center, Shanghai, People's Republic of China: <sup>5</sup>Mass Spectrometry Center, University of Washington, Seattle, WA, USA: <sup>6</sup>Department of Chemistry, University of Helsinki, Helsinki, Finland

Submitted 2 December 2003; Accepted 12 February 2004

### Abstract

**Objective:** The food-frequency questionnaire (FFQ) can be an efficient tool to evaluate dietary intake in large, population-based studies, especially for specific foods. The objective of this study was to validate the assessment of soy and isoflavone (daidzein and genistein) intakes, measured by an FFQ, by comparing intakes with serum isoflavone concentrations.

**Design and setting:** Soy and isoflavone intakes and serum isoflavone concentrations were determined as part of a case–control study of dietary factors and risks of benign breast disease and breast cancer. The FFQ, administered during an in-person interview, included six soy-specific line items. Blood was drawn within one week of FFQ completion.

**Subjects:** In total, 1823 women living in Shanghai, People's Republic of China.

**Results:** In this population, soybean milk, fresh bean curd and other bean foods were eaten once per week, and fermented bean curd, fried bean curd puff and soybeans were eaten less than once per week. A significant linear trend ( $P < 0.01$ ) in serum isoflavone concentrations across increasing categories of soy and isoflavone intakes was observed, indicating that soy and isoflavone intakes, measured by the FFQ, well distinguished serum isoflavone concentrations. Linear trends were also observed in both case and control groups in stratified analyses, suggesting little differential bias by case–control status.

**Conclusions:** The results suggest that the FFQ provides a useful marker of soy food consumption and isoflavone exposure in this population.

### Keywords

Soy  
Food-frequency questionnaire  
Isoflavone  
Serum

Ecological studies suggest that the risk of certain cancers, particularly hormone-associated cancers such as breast cancer, is lower in high soy-consuming populations than low soy-consuming populations<sup>1</sup>. Soy contains isoflavones, which exert a variety of biological activities in animal models and *in vitro*, including anti-oestrogenic, anti-angiogenic, antiproliferative, antioxidant and weak oestrogenic effects<sup>2–4</sup> – effects which may be important in tumour initiation and progression.

In addition to, or apart from, the physiological effects of soy components, soy intake may be a marker of other dietary and lifestyle behaviours that influence disease risk. Within a diet, soy may be consumed in the place of other foods, such as high-fat meats. In Western countries, people who are soy consumers may also engage in other healthy

behaviours, including greater exercise frequency and higher fruit and vegetable consumption<sup>5,6</sup>.

It is difficult to measure associations between soy intake and breast diseases in populations in which soy intake is homogeneously high or low, and therefore studies in populations that have sufficient variation in intake are needed to detect associations between soy intake and disease. Several studies have observed considerable variation in soy food intake in Chinese populations<sup>7–9</sup>. Because the food-frequency questionnaire (FFQ) is an efficient tool to evaluate dietary intake in large, population-based studies<sup>10</sup>, it would be advantageous to measure soy intake by means of an FFQ. The objective of the present analysis was to evaluate the validity of soy intake measured by an FFQ as a surrogate for isoflavone

\*Corresponding author. Email jlampe@fhcrc.org

exposure, by comparing reported intake with serum isoflavones in a population of women living in Shanghai, People's Republic of China.

## Materials and methods

### Participants and study procedures

Participants ( $n = 2315$ ) were drawn from women who participated in a case-control study of diet and breast diseases nested within a randomised trial of breast self-examination among textile workers in Shanghai, China ( $n = 266\,064$ ) (Fig. 1). Recruitment into the trial took place between 1989 and 1991, and follow-up continued to 2000. Women born between 1925 and 1958 were eligible to participate in the trial (aged 30–64 years at enrolment). Further details regarding eligibility for the randomised trial and baseline data collection have been published elsewhere<sup>11</sup>. Dietary data were collected from women in the cohort with benign or malignant breast disease diagnosed between 1 September 1995 and 31 July 2000 and their frequency age-matched controls (Nelson ZC, *et al.*, in preparation). Case women were eligible for the case-control study of diet and breast diseases if they were diagnosed at one of the three hospitals serving the Shanghai Textile Industry Bureau (STIB).

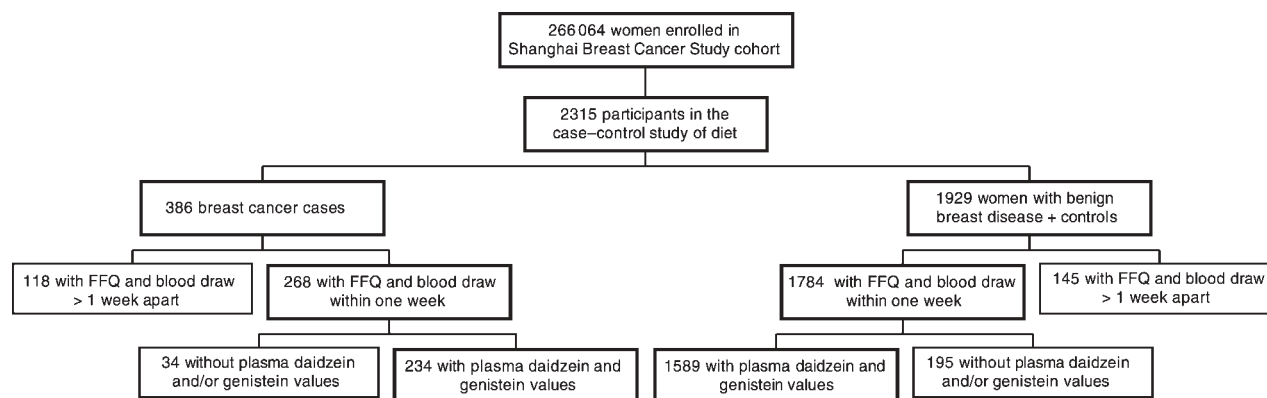
Two non-fasting blood samples were collected for the analysis of several nutrients, including isoflavones. Attempts were made to schedule the diet interview and blood draw on or prior to the date of the scheduled breast biopsy for women with benign or malignant breast disease. Women were excluded from the present analysis if their blood draw was completed and their FFQ was administered greater than one week apart ( $n = 263$ ) or if serum daidzein or genistein was not able to be determined ( $n = 229$ ). The latter resulted primarily from limitations of the assays (more information provided in Results section). Reasons for not having blood drawn and FFQ completed within the same week included good evidence of malignancy at the time of initial breast lump evaluation;

if this was the situation, women were admitted immediately to the hospital for surgery and their blood was drawn after completion of the initial treatment course. Also, if women did not wish to provide two blood samples at the same time for personal or cultural reasons, the second blood sample was scheduled for a month later, so that, for menstruating women, the two specimens were collected in the same phase of the menstrual cycle. Results are presented for 1823 women. The Institutional Review Boards at the Fred Hutchinson Cancer Research Center (FHCRC) and STIB approved the study, and informed consent by all subjects was obtained prior to their participation.

### FFQ and isoflavone content of foods

The FFQ was administered during an in-person interview. The FFQ consisted of 107 food items, including six soy items. The six soy items were: (1) soybean milk, (2) fried bean curd puff, (3) fresh bean curd (tofu), (4) soybeans, (5) fermented bean curd and (6) other bean foods like low-water-content bean curd, soybean curd sheet and vegetarian chicken. Women were asked 'During most of your adult life how often did you usually eat...' for each of the FFQ food items. Women responded as number of times per day, week, month or year, or indicated that they have never consumed the item. Women were also asked whether each item (yes/no) was consumed yesterday.

To calculate dietary isoflavone intake, median serving size for the population was based on publicly available data from the China Health and Nutrition Survey<sup>12</sup>. Values for daidzein and genistein for each FFQ line item were determined using the US Department of Agriculture (USDA)–Iowa State University (ISU) Isoflavones Database<sup>13</sup> and data in Franke *et al.*<sup>14</sup>. The USDA–ISU Isoflavones Database<sup>13</sup> is a collaborative effort between the Food Composition Laboratory and the Nutrient Data Laboratory of the Agricultural Research Service/USDA and the Department of Food Science and Human Nutrition of the ISU. Data for isoflavone content of foods were



**Fig. 1** Participants included in validity analysis of the isoflavone food-frequency questionnaire (FFQ), from the 2315 female Shanghai textile workers who participated in a diet study nested within a randomised trial of breast self-examination. Darker outlined boxes indicate participants included in analysis ( $n = 1823$ )

collected from scientific articles published in refereed journals. Isoflavone data were also obtained by sampling of soy-containing foods and subsequent analysis at ISU. The mean values for daidzein and genistein and the references utilised are presented in the database. Since publication of this database, Franke *et al.*<sup>14</sup> reported analyses of the isoflavone content of soy-containing foods consumed by populations in Singapore and Hawaii. This publication includes analyses of most of the soy items listed in the Shanghai FFQ. The summary values for daidzein and genistein for each line of the Shanghai FFQ were calculated by computing a weighted average of the values in the USDA–ISU database and Franke *et al.*'s values. The values from the USDA–ISU database were weighted by the number of references listed for the food item. As an example for fermented bean curd, the USDA–ISU database lists 143 mg daidzein per kg in 'soybean curd, fermented' (one reference), 143 mg kg<sup>-1</sup> in 'tofu, fermented' (two references), and Franke *et al.* measured the daidzein content to be 250 mg kg<sup>-1</sup>. Therefore, the summary daidzein value for fermented bean curd is  $[(143 \times 1) + (143 \times 2) + (250 \times 1)]/4 = 170 \text{ mg kg}^{-1}$ . Overall, the weight represents the number of articles cited for each food item; thus, each article providing information for a particular food item is given equal weight. Values for fresh bean curd and soybeans were limited to cooked items. Table 1 presents a summary of isoflavone values and calculated isoflavone content per serving for each of the soy foods in the Shanghai FFQ.

#### Serum daidzein and genistein analyses

Blood specimens were collected into 10-ml Vacutainer tubes without anticoagulant. After the blood draw, the tubes were allowed to stand upright at room temperature for at least 30, but not more than 45, min following collection in order to allow the contents to clot. If not processed immediately at 45 min, the tube was placed upright in a cooler with blue ice or placed in a refrigerator until processing. Blood samples were processed into serum within 2 h of blood draw. The sample was centrifuged at 1300 g for 10 min. Serum was aliquoted into 0.5-ml cryovials, and frozen at -20°C for 1 month or at -70°C for longer periods, until being shipped to Seattle.

Specimens were shipped to Seattle on dry ice and immediately stored at -70°C upon arrival at FHCRC.

Serum isoflavone analysis was conducted using two different methods. The first 462 samples (25%) were analysed using liquid chromatography–coulometric array detection (LC–coularray). To improve assay efficiency and precision of the serum daidzein and genistein measurements and given greater access to necessary instrumentation, the remaining 1361 samples (75%) were analysed using liquid chromatography–mass spectrometry (LC–MS). Samples were batched with equal numbers of cases and controls in each batch.

For both LC–coularray and LS–MS, serum concentrations less than 3.9 nmol l<sup>-1</sup> for daidzein and 3.7 nmol l<sup>-1</sup> for genistein were considered below detectable limit (BDL) and were assigned the midpoint value of 2.0 nmol l<sup>-1</sup> for daidzein and 1.9 nmol l<sup>-1</sup> for genistein. 'Undetectable' represents values measured as lower than the limit of detection and is distinct from values that were 'unable to be determined' (as discussed in Participants and study procedures); values 'unable to be determined' were excluded from analysis.

#### LC–coularray

Serum samples were frozen and stored at -70°C until analysis of serum concentrations of daidzein and genistein by LC–coularray using a modified version of the method of Gamache and Acworth<sup>15</sup>. Briefly, 1 ml of serum was hydrolysed enzymatically and extracted with ether. Oestriol glucuronide was added as internal standard before hydrolysis to monitor for loss. High-performance liquid chromatography (HPLC) analysis was performed using a 1050 series pump and autosampler (Agilent Technologies, Palo Alto, CA, USA) coupled to an 8-channel CoulArray 5600 detector (ESA, Inc., Chelmsford, MA, USA) using CoulArray for Windows software. HPLC column, mobile phase, gradient and detector settings were the same as those described previously<sup>15</sup>. Detection limits were 3.9 nmol l<sup>-1</sup> for daidzein and 3.7 nmol l<sup>-1</sup> for genistein. Two serum controls were run with every batch of 20 study samples. Inter-day coefficient of variation (CV) for daidzein (mean = 75.6 nmol l<sup>-1</sup>) and genistein (mean = 72.6 nmol l<sup>-1</sup>) was 17% and 22%, respectively.

**Table 1** Serving size, summary isoflavone values and calculated isoflavone content per serving for each of the soy foods in the Shanghai FFQ

Food item	Serving size (g)	Daidzein (mg kg <sup>-1</sup> )	Genistein (mg kg <sup>-1</sup> )	Daidzein per serving (mg)	Genistein per serving (mg)
Fermented bean curd	55	170	224	9.35	12.32
Soy bean milk	50	46	61	2.30	3.05
Fried bean curd puff	50	141	280	7.05	14.00
Fresh bean curd	75	120	136	9.00	10.20
Other bean foods	50	158	227	7.90	11.35
Soybeans	50	352	328	17.60	16.40

FFQ – food-frequency questionnaire.

**LC-MS**

For analysis by LC-MS, serum samples, frozen and stored at  $-70^{\circ}\text{C}$  until analysis, were extracted and analysed using a modification of the procedures of Lundh *et al.*<sup>16</sup> and Coward *et al.*<sup>17</sup>. To serum samples (0.5 ml), we added an internal standard solution of 10 ng each of  $d_4$ -genistein and  $d_3$ -daidzein (prepared by Dr Kristiina Wähälä, University of Helsinki, Helsinki, Finland) and 0.98 ml of 0.1 M acetate buffer (pH 5) containing 20 U ml<sup>-1</sup> glucuronidase (Boehringer Mannheim, Indianapolis, IN, USA) and 2.0 U ml<sup>-1</sup> sulfatase (Sigma, St Louis, MO, USA). Samples were incubated overnight at 37°C and then extracted using a STRATA C18E normal phase (silica 200 mg/3.0 ml) solid-phase extraction column (Phenomenex, Torrance, CA, USA). The eluent was removed under nitrogen at 40°C and the residue was reconstituted in 100 µl methanol.

LC-MS was performed on the extracts using a series 1100 LC/MSD instrument and a ZORBAX Extend 2.1 mm × 50 mm C-18 column (Agilent Technologies).

The column effluent was subjected to negative-ion electrospray ionisation using nitrogen as a drying gas at a temperature of 350°C and a flow rate of 10.01 min<sup>-1</sup>, with a nebuliser pressure of 20 psi and a capillary voltage of 3.5 kV. Selected ion monitoring data acquisition and processing were accomplished using ChemStation software (Hewlett Packard, Palo Alto, CA, USA) operating in the high-resolution mode. Inter-day variability over four months of analysis was 6–7% CV for daidzein (mean = 98.8 nmol l<sup>-1</sup>) and 4–5% for genistein (mean = 129 nmol l<sup>-1</sup>).

**Statistical methods**

Summary statistics of geometric means and 95% confidence intervals, medians and ranges were calculated. Serum concentration data were log-transformed prior to analyses. Soy food frequency and isoflavone intake were divided into quartiles for analysis. Pearson correlations were calculated between serum daidzein and genistein concentrations. Linear regression was used to evaluate differences in mean serum daidzein and genistein

concentrations across the two assay methods and to analyse trends in serum concentrations across quartiles. Because dietary intake was measured at the time of disease diagnosis and the presence of disease might influence dietary intake, recall of diet or serum markers of diet, comparison of dietary intake and serum isoflavone concentrations was also stratified by case-control status to examine for differential bias. Interactions between case-control status and soy/isoflavone intake and between soy consumption the previous day and soy/isoflavone intake were also modelled using linear regression. Because some analyses in the parent study may be adjusted for energy intake, additional analyses were performed, adjusting for total energy intake. Because season may influence serum isoflavone concentrations, additional analyses, adjusting for season in which the serum measurement was taken, were performed. Season was defined as four 3-month categories; categories were based on average temperature<sup>18</sup>. All analyses were adjusted for method of isoflavone analysis (LC-MS vs. LC-couarray). Statistical significance was considered at  $P < 0.05$ . Analyses were conducted using Stata 8.0 (Stata Corp., College Station, TX, USA).

**Results**

Intake summaries for each of the soy line items in the Shanghai FFQ and total soy intake calculated from these items are presented in Table 2. Two participants reported never consuming any of the soy line items. Generally, median intake of individual soy foods was low, with soybean milk, fresh bean curd and other bean foods each eaten once per week and fermented bean curd, fried bean curd puff and soybeans each eaten less than once per week. At least 70% of women reported consuming, in their adult life, at least one of the soy items. Approximately half of the women had eaten soy the previous day, with soybean milk and other bean foods being the most frequently consumed. Most women reported that, relative to five years ago, they were eating the same amount of soy ( $n = 869$ , 48%) or more soy ( $n = 786$ , 43%) rather than less soy ( $n = 169$ , 9%).

**Table 2** In 1823 female Shanghai textile workers, the median and range of weekly consumption frequency of each soy line item in the Shanghai FFQ, the frequency of ever/never consumption in adult life, whether the item was eaten yesterday, and the average weekly isoflavone intake contributed by each food item

Item	Median (weekly frequency)	Range (weekly frequency)	Ate in adult life, <i>n</i> (%)	Ate yesterday, <i>n</i> (%)	Mean* daidzein (mg)	Mean* genistein (mg)
Fermented bean curd	0.1	0–14	1106 (61)	96 (5)	1.3	1.5
Soybean milk	1	0–35	1331 (73)	320 (18)	2.4	2.9
Fresh bean curd	1	0–7	1812 (99)	261 (14)	9.3	10.5
Fried bean curd puff	0.3	0–14	1612 (88)	85 (5)	1.9	3.4
Soybeans	0.1	0–7	1302 (71)	36 (2)	1.2	1.2
Other bean foods†	1	0–7	1808 (99)	506 (28)	7.4	10.5
Total	4.5	0–43	1824 (100)	971 (53)	29.3	39.3

FFQ – food-frequency questionnaire.

\* Geometric mean; mean intake of daidzein/genistein contributed from individual and total soy food items.

† Examples provided on the FFQ for other bean foods: low-water-content bean curd, soybean curd sheet and vegetarian chicken.

Overall, geometric mean serum daidzein concentration was 49.6 nmol l<sup>-1</sup> (range: BDL–1610 nmol l<sup>-1</sup>) and geometric mean serum genistein concentration was 87.0 nmol l<sup>-1</sup> (range: BDL–3550 nmol l<sup>-1</sup>). Ninety-four women (5%) had undetectable serum daidzein concentrations (2% of samples analysed by LC–coularray and 6% of samples analysed by LC–MS). Fifty-one women (3%) had undetectable serum genistein concentrations (4% of the samples analysed by LC–coularray and 2% of samples analysed by LC–MS). Mean serum daidzein concentrations were higher ( $P < 0.01$ ) in samples analysed by LC–coularray (geometric mean = 76.0 nmol l<sup>-1</sup>, range: BDL–1610 nmol l<sup>-1</sup>) than in samples analysed by LC–MS (geometric mean = 42.9 nmol l<sup>-1</sup>, range: BDL–1440 nmol l<sup>-1</sup>). Mean serum genistein concentrations were similar by analysis method ( $P = 0.33$ ): LC–coularray (geometric mean = 92.2 nmol l<sup>-1</sup>, range: BDL–1930 nmol l<sup>-1</sup>) and LC–MS (geometric mean = 85.5 nmol l<sup>-1</sup>, range: BDL–3550 nmol l<sup>-1</sup>). Correlation between serum daidzein and genistein was 0.84 ( $P < 0.01$ ). Correlations between daidzein and genistein were higher for samples measured by LC–MS ( $r = 0.89$ ,  $P < 0.01$ ) than by LC–coularray ( $r = 0.69$ ,  $P < 0.01$ ). Of the 229 individuals for whom isoflavone concentrations were not able to be determined, 191 (83%) were analysed by LC–coularray. This analytical approach is challenged by non-fasting blood samples with low isoflavone concentrations. Due to interfering peaks on the chromatogram, daidzein and genistein could not be confirmed and guaranteed by coularray for 148 and 100 samples, respectively.

Among all women, significant linear trends ( $P < 0.01$ ) were observed across increasing categories of soy food frequency and across categories of isoflavone intake in geometric mean serum daidzein (Table 3) and genistein (Table 4) concentrations. There was approximately a twofold increase in serum daidzein and genistein concentrations from the lowest to the highest quartiles.

No differences were observed between unadjusted analysis and analysis adjusted for isoflavone assay method (unadjusted data not shown). Linear trends were somewhat stronger for controls than for cases, but no significant interaction was observed between soy intake and case–control status with serum daidzein ( $P$ -value for interaction  $\geq 0.27$ ) or genistein ( $P$ -value for interaction  $\geq 0.11$ ) concentrations. Mean serum daidzein and genistein concentrations were higher for women who consumed soy the previous day than for women who did not consume soy the previous day, but significant linear trends were observed in both groups of women. No significant interaction between previous-day soy consumers and non-consumers with serum daidzein or genistein was observed. Overall results were similar for serum daidzein and serum genistein, but increasing trends appeared to be stronger for serum genistein than for daidzein.

Restricting analysis to samples analysed by LC–MS did not appreciably change or alter the overall interpretation of the results (data not shown). Adjusting for total energy intake or adjusting for season did not appreciably alter these results. For both analyses adjusted for energy or adjusted for season, significant linear trends ( $P < 0.01$ ) were observed for soy frequency intake and isoflavone intake with serum isoflavone concentrations. After adjustment for energy intake or season, similar adjusted values and confidence intervals were observed for analyses stratified on case–control status and on previous-day soy consumption (data not shown).

## Discussion

We observed increasing trends of mean serum daidzein and genistein concentrations across increasing categories of soy food frequency and genistein and daidzein intakes. Because genistein and daidzein are the predominant isoflavonoids in soy foods, these compounds provide

**Table 3** Mean\* serum concentrations of daidzein (nmol l<sup>-1</sup>) by soy food servings and daidzein intake in 1823 female Shanghai textile workers: for all women, stratified by case–control status and stratified by soy consumption the previous day

Soy food intake	All women ( <i>n</i> = 1823)	Controls ( <i>n</i> = 1590)	Cases ( <i>n</i> = 233)	Did not eat soy yesterday ( <i>n</i> = 852)	Ate soy yesterday ( <i>n</i> = 971)
<b>Soy frequency (servings per week)</b>					
Q1: 0–2.7 ( <i>n</i> = 450)	39.0 (34.6, 44.1)	39.8 (34.6, 45.3)	34.2 (23.2, 50.4)	29.9 (25.6, 34.6)	63.0 (51.6, 76.8)
Q2: 2.7–4.5 ( <i>n</i> = 444)	42.1 (37.0, 48.0)	42.5 (37.4, 48.8)	38.2 (26.0, 55.9)	29.5 (25.2, 34.6)	64.2 (53.5, 76.8)
Q3: 4.5–7.7 ( <i>n</i> = 472)	50.4 (44.1, 56.7)	50.8 (44.1, 57.9)	47.6 (33.9, 66.1)	31.5 (26.8, 37.8)	75.2 (63.4, 88.2)
Q4: 7.7–43.0 ( <i>n</i> = 457)	72.8 (64.2, 83.1)	78.3 (68.1, 90.2)	51.6 (37.4, 70.9)	40.6 (31.1, 52.0)	87.4 (76.0, 100)
<i>P</i> for trend†	<0.01	<0.01	0.09	0.10	<0.01
<i>P</i> for interaction†	N/A		0.40		0.45
<b>Daidzein intake (mg week<sup>-1</sup>)</b>					
Q1: 0–20.5 ( <i>n</i> = 455)	40.6 (35.8, 46.1)	40.9 (35.4, 46.5)	39.0 (26.8, 56.7)	29.5 (25.2, 34.3)	68.1 (56.3, 83.1)
Q2: 20.5–30.0 ( <i>n</i> = 456)	43.7 (38.6, 50.0)	44.9 (39.4, 51.5)	35.4 (23.6, 35.5)	30.7 (26.0, 35.5)	63.8 (53.9, 75.6)
Q3: 30.0–43.0 ( <i>n</i> = 456)	50.8 (44.9, 57.8)	53.1 (46.5, 61.0)	35.8 (24.8, 52.4)	31.1 (26.0, 37.8)	72.4 (61.4, 84.6)
Q4: 43.0–206.6 ( <i>n</i> = 457)	67.3 (59.8, 76.8)	70.1 (61.0, 80.7)	57.5 (47.5, 76.8)	37.0 (29.5, 46.1)	88.2 (76.8, 102)
<i>P</i> for trend†	<0.01	<0.01	0.09	0.21	<0.01
<i>P</i> for interaction†	N/A		0.61		0.53

Controls – women without any disease and women with benign breast disease; cases – women with breast cancer; N/A – not applicable.

\* Means presented as geometric mean (95% confidence interval).

† Adjusted for analysis method of the serum isoflavones.



**Table 4** Mean\* serum concentrations of genistein (nmol l<sup>-1</sup>) by soy food servings and genistein intake in 1823 female Shanghai textile workers: for all women, stratified by case-control status and stratified by soy consumption the previous day

Soy food intake	All women (n = 1823)	Controls (n = 1590)	Cases (n = 233)	Did not eat soy yesterday (n = 852)	Ate soy yesterday (n = 971)
Soy frequency (servings per week)					
Q1: 0–2.7 (n = 450)	65.2 (56.7, 74.4)	65.2 (56.7, 75.9)	60.7 (39.6, 91.8)	47.0 (39.6, 55.2)	117 (93.7, 146)
Q2: 2.7–4.5 (n = 444)	71.5 (62.2, 81.5)	71.5 (62.2, 82.9)	67.4 (44.1, 101)	46.7 (38.8, 55.5)	118 (97.4, 144)
Q3: 4.5–7.7 (n = 472)	92.6 (80.4, 106)	93.7 (81.5, 108)	88.1 (61.5, 128)	57.8 (47.8, 69.2)	140 (117, 166)
Q4: 7.7–43.0 (n = 457)	133 (117, 153)	144 (124, 167)	90.0 (64.1, 128)	70.7 (54.1, 92.6)	161 (138, 187)
P for trend†	<0.01	<0.01	0.11	0.01	<0.01
P for interaction†	N/A		0.21		0.87
Genistein intake (mg week <sup>-1</sup> )					
Q1: 0–26.8 (n = 456)	65.2 (56.7, 75.2)	65.9 (56.7, 76.7)	61.5 (41.1, 92.6)	45.2 (38.1, 53.3)	123 (98.5, 151)
Q2: 26.8–39.0 (n = 456)	76.7 (66.7, 87.4)	78.1 (67.4, 90.0)	62.2 (39.6, 97.4)	51.5 (42.6, 61.5)	114 (95.5, 136)
Q3: 39.0–57.0 (n = 456)	95.5 (83.7, 110)	98.5 (84.8, 113)	82.2 (56.2, 119)	56.2 (45.9, 68.1)	146 (123, 174)
Q4: 57.0–262.4 (n = 456)	121 (106, 138)	126 (109, 147)	97.4 (70.0, 136)	64.1 (50.0, 81.5)	159 (136, 185)
P for trend†	<0.01	<0.01	0.06	0.03	<0.01
P for interaction†	N/A		0.56		0.95

Controls – women without any disease and women with benign breast disease; cases – women with breast cancer; N/A – not applicable.

\* Means presented as geometric mean (95% confidence interval).

† Adjusted for analysis method of the serum isoflavones.

more suitable biomarkers of soy intake than other soy isoflavonoid compounds such as glycitein and biochanin A. We observed trends that, with the exception of daidzein intake in cases, appeared to increase in a linear manner, suggesting that increasing intake distinguished serum genistein and daidzein concentrations. Small differences in the linear trends for cases and controls were not statistically significant. It is unlikely that the smaller sample size of cases explains the lack of statistical significance in the interaction between cases and controls; non-significance most likely reflects the fact that, although the magnitude was somewhat smaller for cases, the direction of association was the same in cases and controls, suggesting no substantial differential bias in soy intake classification by case-control status. Linear trends in mean serum daidzein and genistein concentrations were observed across quartiles of soy food frequency and isoflavone intake, regardless of whether soy was consumed the previous day or not, suggesting that the FFQ captures soy intake variation beyond most recent soy consumption. Overall, the results we observed suggest that soy food frequency determined by the FFQ provides a valid measure of soy intake and isoflavone exposure.

Other studies comparing soy or isoflavone intake assessed by FFQ or dietary records with overnight urine or plasma isoflavone concentrations, which evaluated intake on a continuous scale, reported correlation coefficients ranging from 0.31 to 0.54<sup>8,19–22</sup>. Our analyses focused on categorised intake of isoflavones, rather than continuous intake, for several reasons. First, intake and serum values of these compounds did not follow a normal distribution whether log-transformed or not; thus, correlation analyses may be strongly influenced by small numbers of individuals with high values for intake or serum concentrations. Second, because we did not collect

information on serving size, we would expect that continuous intake may not provide a wholly accurate estimate of an individual's intake; thus, categories may provide a more reliable grouping of individuals by intake. Third, analyses with disease outcome in the parent study will be conducted using quartiles of intake. Willett<sup>23</sup> discusses several advantages for this approach, including visualisation of how actual numbers of cases and controls vary by level of intake, visualisation of trends as being compatible with linearity or non-linearity, and constraining any undue influence of outlying data points which may be most affected by errors in data collection.

Few large-scale observational studies have measured serum isoflavone concentrations. Average concentrations observed in several Japanese studies range from 120 to 283 nmol l<sup>-1</sup> for daidzein and from 415 to 492 nmol l<sup>-1</sup> for genistein<sup>19,24,25</sup>. In the USA, we observed considerably lower average serum isoflavone concentrations in previous studies: 6.89–8.90 nmol l<sup>-1</sup> for daidzein and 12.1–16.9 nmol l<sup>-1</sup> for genistein<sup>26,27</sup>. Similar to the US studies, Morton *et al.*<sup>25</sup> observed average serum isoflavone values in British men of 17.9 nmol l<sup>-1</sup> for daidzein and 33.2 nmol l<sup>-1</sup> for genistein. In the present study we observed serum concentrations of daidzein and genistein of 50.0 nmol l<sup>-1</sup> and 87.0 nmol l<sup>-1</sup>, respectively, suggesting that isoflavone exposure in this population of Shanghai women is less than that of Japanese populations, but higher than that of Western populations.

Soy intake, as measured by weekly consumption of isoflavones, appears to be lower than in other Chinese populations. We observed that women in this study consumed, on average, 29 mg of daidzein and 39 mg of genistein per week. This is within the range of isoflavone intake that has been observed in Western populations (<1 to 80 mg week<sup>-1</sup>)<sup>21,28</sup> but is less than that observed in other studies of urban Chinese women, which have

observed average isoflavone intakes in the range of 82 to 286 mg per week<sup>7,9,29,30</sup>, including one study conducted in Shanghai, in which weekly mean isoflavone intake was 286 mg<sup>7</sup>. Lower intake in this population may indicate a shift in diet towards less soy food consumption. However, only 9% of women reported eating less soy now than five years previously. The lower estimated isoflavone intake may also be a consequence of using the median serving size of the population, which may have been underestimated, for all participants in our study.

There are several potential limitations in this study related to the collection of dietary data and to the measurement of serum isoflavones. One limitation was that median serving size for the population was used to calculate isoflavone intake for all participants. Willett summarises results from several studies that indicate that individuals are relatively inaccurate in recalling portion sizes and that the majority of intake variability is from frequency of consumption rather than serving size<sup>23</sup>. Thus, not collecting serving size information likely reduced the intake values but did not markedly alter an overall ranking of individuals by intake. However, if reduction in the variability in dietary intake did occur, this would represent a source of non-differential bias and would have reduced the ability to detect an association between soy intake and serum isoflavone concentrations. In this circumstance, our results would be conservative.

With respect to measurement of serum isoflavones, one limitation is the difference in mean serum daidzein concentrations across isoflavone methods. This was not observed for genistein and was not observed for quality control samples tested by both assays. There are several possible reasons for the discrepancy. With consumption of soy foods, serum daidzein concentrations are 40% lower than genistein concentrations, so are more likely to be near the detection limit with low soy intake. As illustrated in Table 1, compared with genistein, daidzein is typically found in lower concentrations in food and this likely contributes to lower daidzein concentrations in serum. Given the chromatography conditions of our assay, daidzein also eluted at a time where there tended to be more peaks, compared with genistein. Both of these factors made confirmation more difficult for daidzein than for genistein. None the less, the relationships between dietary intake and serum concentrations were similar for genistein and daidzein, suggesting that this difference across assay methods contributed little to the overall findings. Another concern is that we were unable to confirm daidzein (46% of unconfirmed samples), genistein (24%) or both (30%) for 30% of the samples analysed by LC–coularray (9% of samples overall). The limitations of this assay resulted primarily in the exclusion of samples with low isoflavone concentrations, which would represent a differential bias of exclusion by soy intake group. Although it is uncertain whether and in which direction these exclusions biased the results, we hypothesise that

these exclusions would be likely to inflate the mean serum isoflavone values in the smallest intake category, and thus would reduce the slope of the association across intake categories and attenuate the overall association. None the less, in our study, adjustment for analysis method did not change the association between dietary intake and serum concentrations of daidzein or genistein, suggesting that these differences did not markedly influence our results. Another limitation is that blood samples were non-fasting, which would also probably result in non-differential bias in serum concentrations in relation to soy intake. Because a single serum sample was collected rather than multiple samples over time and that single serum sample was compared with usual adult intake, we would expect decreased ability to detect an association. A final limitation is that because the half-life is approximately 4–8 h for daidzein and 6–8 h for genistein<sup>31–33</sup>, serum concentrations are generally considered a biomarker for recent intake. Urinary isoflavone measures are considered to integrate exposure over a longer time period; however, blood and urinary isoflavones in relation to dietary intake are similarly correlated<sup>20</sup>, suggesting that serum concentrations provide a similar ranking in relation to dietary intake. Given this short half-life of serum isoflavones, repeated blood isoflavone measurements would likely improve the associations between intake and serum concentrations observed in this study. Overall, there are several sources of potential non-differential bias in our study; thus, the associations between serum isoflavones and dietary soy and isoflavone intake we observed may be conservative.

Despite a number of sources of potential non-differential bias, we observed a linear trend across serum isoflavone concentrations, objective biomarkers of soy intake<sup>34</sup>, with increasing soy food consumption frequency and with increasing isoflavone intakes, subjective markers of soy intake. These results suggest that soy intake measured by the FFQ provides a useful marker of soy intake and isoflavone exposure in this population of women.

## Acknowledgements

This work was supported by NIH R01CA46823. C.L.F. was also supported, in part, by NCI T32CA009168.

## References

- 1 Adlercreutz H, Mazur W. Phyto-oestrogens and Western diseases. *Annals of Medicine* 1997; **29**: 95–120.
- 2 Hopert A-C, Beyer A, Frank K, Strunck E, Wünsche W, Vollmer G. Characterization of estrogenicity of phytoestrogens in an endometrial-derived experimental model. *Environmental Health Perspectives* 1998; **106**: 581–6.
- 3 Breinholt V, Hossaini A, Svendsen GW, Brouwer C, Nielson SE. Estrogenic activity of flavonoids in mice. The importance of estrogen receptor distribution, metabolism and bioavailability. *Food and Chemical Toxicology* 2000; **38**: 555–64.

- 4 Setchell KDR. Phytoestrogens: the biochemistry, physiology, and implications for human health of soy isoflavones. *American Journal of Clinical Nutrition* 1998; **68**(Suppl.): 1333S–46S.
- 5 Guthrie JR, Ball M, Murkies A, Dennerstein L. Dietary phytoestrogen intake in mid-life Australian-born women: relationship to health variables. *Climacteric* 2000; **3**: 254–61.
- 6 Keinan-Boker L, Peeters PHM, Mulligan AA, Navarro C, Slimani N, Mattison I, *et al.* Soy product consumption in 10 European countries: the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Public Health Nutrition* 2002; **5**: 1217–26.
- 7 Dai Q, Shu XO, Jin F, Potter JD, Kushi LH, Teas J, *et al.* Population-based case–control study of soyfood intake and breast cancer risk in Shanghai. *British Journal of Cancer* 2001; **85**: 372–8.
- 8 Chen Z, Zheng W, Custer LJ, Dai Q, Shu XO, Jin F, *et al.* Usual dietary consumption of soy foods and its correlation with the excretion rate of isoflavonoids in overnight urine samples among Chinese women in Shanghai. *Nutrition and Cancer* 1999; **33**: 82–7.
- 9 Ho SC, Woo JL, Leung SS, Sham AL, Lam TH, Janus ED. Intake of soy products is associated with better plasma lipid profiles in the Hong Kong Chinese population. *Journal of Nutrition* 2000; **130**: 2590–3.
- 10 Willett W. *Nutritional Epidemiology*. New York: Oxford University Press, 1990.
- 11 Thomas DB, Gao DL, Ray RM, Wang WW, Allison CJ, Chen FJ, *et al.* Randomized trial of breast self-examination in Shanghai: final results. *Journal of the National Cancer Institute* 2002; **94**: 1445–57.
- 12 Carolina Population Center, University of North Carolina at Chapel Hill. *The China Health and Nutrition Survey* [online], 1999. Available at [http://www.cpc.unc.edu/projects/china/china\\_home.html](http://www.cpc.unc.edu/projects/china/china_home.html)
- 13 US Department of Agriculture–Iowa State University Database on the Isoflavone Content of Foods, Release 1.3 – 2002 [online], 2002. Available at <http://www.nal.usda.gov/fnic/foodcomp/Data/isoflav/isoflav.html>
- 14 Franke AA, Hankin JH, Yu MC, Maskarinec G, Low SH, Custer LJ. Isoflavone levels in soy foods consumed by multiethnic populations in Singapore and Hawaii. *Journal of Agricultural and Food Chemistry* 1999; **47**: 977–86.
- 15 Gamache PH, Acworth IN. Analysis of phytoestrogens and polyphenols in plasma, tissue, and urine using HPLC with coulometric array detection. *Proceedings of the Society for Experimental Biology and Medicine* 1998; **217**: 274–80.
- 16 Lundh T, Pettersson H, Kiessling K. Liquid chromatographic determination of the estrogens, daidzein, formononetin, coumestrol, and equol in bovine plasma and urine. *Journal of the Association of Official Analytical Chemists* 1988; **71**: 938–41.
- 17 Coward L, Kirk M, Albin N, Barnes S. Analysis of plasma isoflavones by reversed-phase HPLC–multiple reaction ion monitoring-mass spectrometry. *Clinica Chimica Acta* 1996; **247**: 121–42.
- 18 Yearly climate info – Shanghai Municipality. *Travel China Guide* [online], 2003. Available at <http://www.travelchinaguide.com/climate/shanghai.htm>
- 19 Yamamoto S, Sobue T, Sasaki S, Kobayahi M, Arai Y, Uehara M, *et al.* Validity and reproducibility of a self-administered food-frequency questionnaire to assess isoflavone intake in a Japanese population in comparison with dietary records and blood and urine isoflavones. *Journal of Nutrition* 2001; **131**: 2741–7.
- 20 Arai Y, Uehara M, Sato Y, Kimira M, Eboshida A, Adlercreutz H, *et al.* Comparison of isoflavones among dietary intake, plasma concentration and urinary excretion for accurate estimation of phytoestrogen intake. *Journal of Epidemiology* 2000; **10**: 127–35.
- 21 Maskarinec G, Singh S, Meng L, Franke AA. Dietary soy intake and urinary isoflavone excretion among women from a multiethnic population. *Cancer Epidemiology, Biomarkers & Prevention* 1998; **7**: 613–9.
- 22 Huang MH, Harrison GG, Mohamed MM, Gornbein JA, Henning SM, Go VL, *et al.* Assessing the accuracy of a food frequency questionnaire for estimating usual intake of phytoestrogen. *Nutrition and Cancer* 2000; **37**: 145–54.
- 23 Willett WC. *Nutritional Epidemiology*, 2nd ed. New York: Oxford University Press, 1998.
- 24 Nagata C, Shimizu H, Takami R, Hayashi M, Takeda N, Yasuda K. Soy product intake and serum isoflavonoid and estradiol concentrations in relation to bone mineral density in postmenopausal Japanese women. *Osteoporosis International* 2002; **13**: 200–4.
- 25 Morton MS, Arisaka O, Miyake N, Morgan LD, Evans BA. Phytoestrogen concentrations in serum from Japanese men and women over forty years of age. *Journal of Nutrition* 2002; **132**: 3168–71.
- 26 Frankenfeld CL, Patterson RE, Kalthorn TF, Skor HE, Howald WN, Lampe JW. Validation of a soy food frequency questionnaire with plasma concentrations of isoflavones in US adults. *Journal of the American Dietetic Association* 2002; **102**: 1407–13.
- 27 Frankenfeld CL, Patterson RE, Horner NK, Neuhaus ML, Skor HE, Kalthorn TF, *et al.* Validation of a soy food-frequency questionnaire and evaluation of correlates of plasma isoflavone concentrations in postmenopausal women. *American Journal of Clinical Nutrition* 2003; **77**: 674–80.
- 28 De Kleijn MJ, van der Schouw YT, Wilson PW, Adlercreutz H, Mazur W, Grobbee DE, *et al.* Intake of dietary phytoestrogens is low in postmenopausal women in the United States: The Framingham Study. *Journal of Nutrition* 2001; **131**: 1826–32.
- 29 Jakes RW, Duffy SW, Ng FC, Gao F, Ng EH. Mammographic parenchymal patterns and self-reported soy intake in Singapore Chinese women. *Cancer Epidemiology, Biomarkers & Prevention* 2002; **11**: 608–13.
- 30 Sun CL, Yuan JM, Arakawa K, Low SH, Lee HP, Yu MC. Dietary soy and increased risk of bladder cancer: the Singapore Chinese Health Study. *Cancer Epidemiology, Biomarkers & Prevention* 2002; **11**: 1674–7.
- 31 Setchell KDR, Faughnan MS, Avades T, Zimmer-Nechemias L, Brown NM, Wolfe BE, *et al.* Comparing the pharmacokinetics of daidzein and genistein with the use of <sup>13</sup>C-labeled tracers in premenopausal women. *American Journal of Clinical Nutrition* 2003; **77**: 411–9.
- 32 Watanabe S, Yamaguchi M, Sobue T, Takahashi T, Miura T, Arai Y, *et al.* Pharmacokinetics of soybean isoflavones in plasma, urine and feces of men after ingestion of 60 g baked soybean powder (kinako). *Journal of Nutrition* 1998; **128**: 1710–5.
- 33 King RA, Bursill DB. Plasma and urinary kinetics of the isoflavones daidzein and genistein after a single soy meal in humans. *American Journal of Clinical Nutrition* 1998; **67**: 867–72.
- 34 Lampe JW. Isoflavonoid and lignan phytoestrogens as dietary biomarkers. *Journal of Nutrition* 2003; **133**(Suppl. 3): 956S–64S.