Reliability of Serum Measurements of Lignans and Isoflavonoid Phytoestrogens over a Two-Year Period

Anne Zeleniuch-Jacquotte, Herman Adlercreutz, Arslan Akhmedkhany, and Paolo Toniolo

Nelson Institute of Environmental Medicine [A. Z-J., A. A., P. T.], Kaplan Comprehensive Cancer Center [A. Z-J., P. T.], and Department of Obstetrics and Gynecology [A. A., P. T.], New York University Medical Center. New York, New York 10010-2598 and Department of Clinical Chemistry, University of Helsinki, and Folkhalsan Research Center, 00280 Helsinki, Finland [H. A.]

Abstract
We examined the distribution and long-term reliability of serum measurements of the two main human lignans, enterolactone and enterodiol, and the isoﬂavonoid phytoestrogens daidzein, genistein, equol, and O-Desmethylangolensin in the New York University Women’s Health Study, a prospective cohort study of sex hormones and breast cancer. Serum samples collected at three yearly visits in 30 premenopausal and 30 postmenopausal women who had not been diagnosed with cancer or cardiovascular disease were included in the study. Assays were carried out by ion-exchange chromatography and capillary gas chromatography-mass spectrometry. Levels of isoﬂavonoid phytoestrogens were low, often at or below the sensitivity level of the assay. The reliability coefﬁcients for these compounds were also low (≤0.30). The median levels of enterodiol and enterolactone were 1.52 nmol/liter and 20.2 nmol/liter, respectively, and were comparable with the levels observed in omnivorous Finnish women living in the Helsinki area. A substantial number of women, though, had fairly high levels: for instance, 15% of the assays showed levels of enterolactone greater than the mean level observed in vegetarian Finnish women, i.e., 89.1 nmol/liter (H. Adlercreutz et al., Cancer Detect. Prev., 18: 259-271, 1994). The reliability coefﬁcient of a single measurement of enterolactone was moderately high (0.55), suggesting that serum measurements of this compound could be a useful tool in prospective epidemiological studies with access to repeated blood or serum specimens. For instance, the reliability coefﬁcient of the average of three measurements of enterolactone would be 0.79, a level considered acceptable in light of the other sources of error that are present in epidemiological studies (W. Willett, Stat. Med., 8: 1031-1040, 1989).

Introduction
Phytoestrogens comprise several classes of chemical compounds including lignans, isoﬂavonoids, and coumestans. They are naturally occurring plant substances found in many foods, and are structurally similar to estradiol but have two phenolic rings. They have numerous biological effects including both estrogen agonist and antagonist properties. Owing to their estrogen antagonist and tumor growth inhibition properties (1), phytoestrogens have been proposed as protective agents against hormone-dependent cancers, in particular breast and prostate cancer (2-5). Their estrogen agonist properties, on the other hand, suggest that phytoestrogens may protect against menopausal symptoms, osteoporosis, and cardiovascular disease (3, 4, 6, 7).

The dietary sources of phytoestrogens are only partially known. Oilseeds, such as flaxseed, have the highest detected contents of lignans. Seaweeds, legumes, cereal brans, whole cereals, vegetables, and fruits contain lesser amounts (8). Mammalian isoﬂavonoids are derived mostly from soy beans and fermented soy products. Some alcoholic beverages, such as beer and bourbon, and several varieties of legumes, such as peas (e.g., chickpeas and green split peas) and beans (e.g., kidney, black-eyed, and lima beans) also contain isoﬂavonoid precursors, but at lower levels (9-11). Unfortunately, the precise phytoestrogen content of many individual foods is not known and may differ according to variety, location, and season, as it has been shown for the isoﬂavonoid content of soybeans (12). As a result, epidemiological studies based on dietary questionnaires are limited in their ability to assess the nutritional intake of phytoestrogens accurately.

Recently, a gas chromatographic-mass spectrometric method has been developed to measure phytoestrogen levels in plasma or serum (13, 14). The availability of this new assay offers the opportunity to examine the associations of phytoestrogens with the risk of chronic diseases in prospective studies with access to frozen blood or serum samples. In this type of epidemiological studies, the exposure of interest is the long-term average blood or serum level, rather than the level at one point in time. It is, therefore, important before undertaking such studies to assess the reliability of serum phytoestrogen measurements (i.e., the consistency or reproducibility of measurements taken at different time points in the same individual).

It is also important to assess whether exposure levels that one is interested in evaluating are represented with sufficient frequency in the study population.

We report here the results of a pilot study to assess the feasibility of a case-control study of phytoestrogens and breast...
cancer nested within the NYU
Women's Health Study, a prospective cohort study of breast cancer and sex hormones and diet. The specific objectives of the pilot study were to estimate the long-term reliability of serum measurements of the two main lignans, enterodiol and enterolactone, and the isoflavonoids daidzein, genistein, equol, and O-Desmethylangolensin, and to examine the distribution of serum phytoestrogen levels in our population. A secondary objective was to assess whether menopausal status, age, and storage time affect serum levels of phytoestrogens.

Materials and Methods
The NYU Women's Health Study. Between March 1985 and June 1991, the NYU Women's Health Study enrolled a cohort of 14,275 healthy women, ages 34–65 years, at a breast cancer screening center in New York City (15, 16). At the time of enrollment and at annual screening visits thereafter, information on medical, anthropometric, and reproductive factors was collected through a self-administered questionnaire, and 30 ml of nonfasting peripheral venous blood was collected. After blood drawing, tubes were kept covered at room temperature (70°F) for 15 min, then at 4°C for 60 min to allow clot retraction, and then centrifuged at 600 × g. Supernatant serum was partitioned into 1-ml aliquots and immediately stored at –80°C for future biochemical analyses. Approximately half of the participants gave blood at repeated visits (mean number of blood donations, 3; range, 2–8). On average, twelve 1-ml serum aliquots were generated/blood donation. Women have since been followed-up for cancer and cardiovascular end points. Nested case-control studies of breast, colon, and endometrial cancer, as well as of coronary heart disease are ongoing.

Sample Selection. NYU Women's Health Study participants who had given blood on three or more occasions with a yield of 11 or more aliquots/visit, who had not been diagnosed with cancer or cardiovascular disease, and who had not been selected as a control in any case-control study nested within the cohort were eligible for the present study. Sampling was stratified by menopausal status, age at enrollment (35–45 years, >45 years for premenopausal, and <55 years, 55–65 years for postmenopausal women), and year of first blood donation (1985–87, 1988–89). Thirty premenopausal and 30 postmenopausal women were selected. Premenopausal women had a mean age (±SD) at first blood donation of 45.7 years (±5.8 years) and a mean Quetelet's index of 22.2 kg/m² (±2.6 kg/m²). For postmenopausal women, the mean age was 57.7 years (±4.8 years) and the mean Quetelet's index was 25.1 kg/m² (±4.5 kg/m²). Mean times in storage of the serum samples were 8.5 years (±1.2 years), 7.5 years (±1.2 years), and 6.5 years (±1.2 years) for visits one, two, and three, respectively.

Because the assay required 2 ml of serum, for each subject, 2 aliquots of serum were retrieved from each of the three selected blood donations. Samples were shipped in 15 batches to the participating laboratory in Finland. Each batch contained 14 samples (28 aliquots): 3 samples (from three blood donations) for each of four subjects plus 2 samples from a standard pool. The standard pool samples were used to assess inter- and intrabatch variability. Because the number of batches was larger than had been initially planned, two different pools were used sequentially, one for the first 10 batches, the other for the last 5 batches. Batches were shipped in dry ice, two at a time, approximately every 2 months. All specimens were identified solely by a code number.

Laboratory Analyses. At the laboratory, the 2 aliquots of each sample were thawed and pooled, to generate the 2 ml necessary for the assays. Serum phytoestrogen concentrations were measured by a modification of a method for determining phytoestrogens in plasma by ion-exchange chromatography and capillary gas chromatography-mass spectrometry. A detailed description of the original method is presented elsewhere (14). Instead of the separation of the sulfate and free fraction from the glucuronide fraction, the sulfates and glucuronides were hydrolyzed in consecutive steps and all free phytoestrogens were extracted with diethyl ether.

Statistical Methods. The measurements on the standard pool aliquots were used to quantify the assay variability. Between 2 aliquots of the same standard pool were included in each batch, we were able to estimate both within- and between-batch variabilities (also called intra- and interassay variabilities) using a variance components model including batch as a random effect. The total laboratory variance was computed as the sum of the intra- and interassay variabilities. Intra- and total assay coefficients of variation were estimated separately for the two pools, and using the raw as well as the log-transformed data. Because observed levels included zero, one was added to all values before log-transformation.

The reliability was estimated by the intraclass correlation coefficient assuming a one-way random effects model ANOVA (17). Estimation was carried out using restricted maximum likelihood techniques in SAS PROC VARCOMP. Computations were performed on the log-transformed data because future epidemiological analyses of serum phytoestrogen levels will most likely be carried out on this scale to reduce the positive skewness of the raw data. Approximate 95% confidence intervals were calculated as described by Donner (17). The potential effect of covariates on phytoestrogen levels was assessed by fitting a general linear model. Variables considered were age, storage time, and menopausal status.

A number of measurements were below 1.0 nmol/liter, the sensitivity level of the assay, defined as the level at which the intra-batch coefficient of variation is 100%. The reliability coefficients presented here were estimated using values as provided by the laboratory. Analyses performed after assigning all values below the sensitivity level to 0.5 nmol/liter, the midpoint between 0 and the sensitivity level, gave similar results and are, therefore, not presented.

Results
Table 1 displays intra- and total assay coefficients of variation for the raw and the log-transformed data. Before log-transformation, coefficients of variation were higher than expected, which most likely resulted from the low phytoestrogen levels in the pools. The coefficients of variation decreased after log-transformation of the data. For matched epidemiological studies, intra-assay coefficients of variation are the most relevant, because samples from a matched set are usually analyzed within the same batch. The intra-assay coefficients of variation on the log-transformed data were low (<5%) for enterolactone, acceptably low for genistein (<10%), but still high for the remaining compounds.

Table 2 reports the phytoestrogen levels observed at each blood sampling. For the lignans, enterodiol values were very low, with the median level around 1.50 nmol/liter at each visit, and the 75th percentile not exceeding 3.24 nmol/liter. Additionally, for over one-third of the serum samples, levels were

1The abbreviations used are: NYU, New York University.
below 1.00 nmol/liter, the sensitivity level of the assay. Levels of enterolactone were substantially higher, with median levels varying from 18.5–22.7 nmol/liter, and only 2% of the assays below 1.00 nmol/liter. Among the isoflavonoids, levels of daidzein were low (median levels were 3.67, 3.54, and 2.11 nmol/liter at visits one, two, and three, respectively), whereas levels of genistein were moderate (median levels were 9.25, 7.04, and 4.86 nmol/liter at visits one, two, and three, respectively). Levels of equol and O-Desmethylangolensin were very low, with approximately two-thirds of the assays below 1.00 nmol/liter.

The original article was published in Cancer Epidemiology, Biomarkers & Prevention (1998).
disease risk. Indeed, they reported substantial heterogeneity in urinary excretion of seven phytoestrogens in a group of 50 Caucasian, African-American, Latina, and Japanese women living in northern California [19].

The reliability coefficients for isoflavonoids were low (0.11–0.30), which may be due, for some of the compounds, to the large proportion of values around the sensitivity level of the assay. Also, high reliability coefficients would be expected for compounds consumed on a regular basis. Although consumption of soy products by our participants was not assessed, it is likely to be occasional because soy products are not an integral part of most Caucasian diets. Low reliability for an exposure measurement will lead to loss of power as well as bias in the observed relative risks. Adequate power may be obtained by increasing the number of subjects, although the increase required may be considerable. For instance, a study where the reliability of the exposure is 0.30 would need a sample size 3.3 times larger than a study where exposure is measured without error [20]. In case of a reliability coefficient of 0.11, the sample size would have to be nine times larger. It should be noted, however, that increments in sample size will not reduce bias in the observed relative risks. Another approach is to improve the accuracy of the exposure measurement by increasing the number of measurements performed on each subject and using the mean of the repeated measurements as the exposure variable. This approach will generally improve both the study power and the accuracy of the estimates of the effect of the exposure. However, the number of measurements required for each subject may also be prohibitive. For instance, 10 measurements would be needed to raise a reliability coefficient from 0.30–0.81, a level considered acceptable in light of the other sources of error that are present in epidemiological studies [21].

The coefficients of variation of the isoflavonoid assays were higher than expected, which may also be a consequence of the low values observed in our group of women. Taken together, the high coefficients of variation, small proportion of high values, and low reliability coefficients of serum levels of isoflavonoids in the NYU Women’s Health Study indicate that the study is not well suited to examine the relation of serum isoflavonoids with disease risk.

The mean serum levels of lignans were comparable with the levels observed in omnivorous Finnish women. A substantial number of women, though, had fairly high levels: for instance, 15% of the assays showed levels of enterolactone greater than the mean level observed in vegetarian Finnish women (i.e., 89.1 nmol/liter). Results were similar for enterodiol, with levels >5.4 nmol/liter for 17% of the assays. The reliability coefficients were higher than for isoflavonoids: 0.37 for enterodiol and 0.55 for enterolactone, respectively. These results indicate that although a single measurement will imperfectly reflect an individual’s average long-term level, serum measurements, in particular of enterolactone, may still be a useful tool in epidemiological studies with access to repeated blood or serum specimens. The reliability of enterolactone would be raised to 0.79 if the average of three measurements was used as exposure variable, whereas seven measurements would be needed to raise the reliability of enterodiol to 0.80. The use of biological specimens in epidemiological studies of lignans is of particular interest because assessment of dietary intake is currently limited by the paucity of quantitative data on the lignan content of individual foods. To our knowledge, no study has yet been published on the relation of dietary intake of lignans with breast cancer risk.

Use of repeated measurements has been advocated for epidemiological studies investigating exposures with known or suspected measurement error [22]. It is usually recommended to use the average of the values observed in repeated samples for each individual. An alternative approach would be to pool the samples of an individual before performing the assay. This method would be of particular interest for expensive assays and/or assays requiring a large volume of serum. Because only one assay/individual would be performed, the cost of the assays would be reduced by a factor of n, the number of repeated samples used. The per-visit volume of biological material needed for the assay would be reduced by the same factor. In studies with prospectively collected biological specimens, the samples of the individuals who are subsequently diagnosed with the disease of interest are very valuable, and strategies to preserve material from these subjects should be given consideration.

The objectives of this pilot study were to estimate the long-term reliability of serum measurements of six lignans or isoflavonoid phytoestrogens, as well as examine their distribution in the NYU Women’s Health Study, a population of middle-aged, Caucasian women living in New York City. We found that levels of isoflavonoids were low, often at or below the sensitivity level of the assay, and associated with low reliability coefficients, indicating that our population is not well suited to study the association of these compounds with disease. On the other hand, for enterolactone, the range of values was quite large and the reliability over a 2-year period was moderately high. These results suggest that the effect of this compound on disease risk could be examined in our study population, especially because repeated serum samples are available for a subset of the participants. The present data on reliability can be used to select the optimum design with respect to sample size and number of samples/participant for future epidemiological studies assessing the association of serum phytoestrogen levels with health effects.

References


Reliability of serum measurements of lignans and isoflavonoid phytoestrogens over a two-year period.


Updated version  Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/7/10/885

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.