Reliability of Serum Prolactin Measurements in Women

Karen L. Koenig, Paolo Toniolo, Peter F. Bruning, J. M. G. Bonfrer, Roy E. Shore, and Bernard S. Pasternack

Abstract
Prolactin, a hormone indispensable for milk secretion, has been shown to enhance the development and growth of mammary tumors in rodents; however, its importance in human breast cancer is uncertain. Serum prolactin levels are known to fluctuate considerably under normal conditions, and lack of precision in the hormone measurements may have contributed to the largely negative findings in humans to date. The purpose of this study was to investigate the reliability of prolactin measurements in women using stored serum from an ongoing prospective study of breast cancer. Separate groups of postmenopausal and premenopausal women who donated multiple blood samples at approximately 1-year intervals were studied. The reliability of a single log prolactin determination, as measured by the intraclass correlation coefficient, was 0.76 for the postmenopausal women (95% confidence interval, 0.66-0.85) and 0.48 for the premenopausal women (95% confidence interval, 0.31-0.62). These findings suggest that a single measurement is sufficient to characterize the serum prolactin level of postmenopausal women for epidemiological research. For premenopausal women, however, multiple samples are desirable. Controlling for phase of the menstrual cycle does not appear to substantially improve the reliability of premenopausal measurements.

Introduction
Prolactin, a hormone indispensable for milk secretion, has been shown to enhance the development and growth of mammary tumors in rodents (1). The importance of prolactin in human breast cancer, however, remains to be demonstrated (2). We are investigating the relation between prolactin levels and subsequent breast cancer risk in the NYU Women's Health Study, a case-control-within-a-cohort study of endogenous hormones and breast cancer (3). An important consideration in this and other epidemiological studies involving prolactin measurements is the reliability of the measurements, since prolactin levels are known to fluctuate considerably under normal conditions. Among the factors affecting prolactin levels are stress, drugs, menopausal status, time in the menstrual cycle, time of day, and time since last meal (4).

The availability of repeat blood samples drawn at approximately 1-year intervals in a large subset of the study cohort gave us the opportunity to study the reproducibility of prolactin measurements in women. The goal of our investigation was to answer the following questions. (a) Is the variability in the level of prolactin within subjects so great that a single measurement reveals little about a woman's long-term exposure? (b) How many measurements are needed to characterize a woman's prolactin exposure status for the purposes of epidemiological research?

Methods
The NYU Women's Health Study. Between March 1985 and June 1991, the NYU Women's Health Study enrolled a cohort of 14,290 women aged 34-65 years at the Guttman Breast Diagnostic Institute, a breast screening clinic in New York City. At the time of enrollment and at annual screening visits thereafter, subjects were asked to complete questionnaires and to provide 30 ml of peripheral venous blood. Fifty-one percent of cohort members donated blood on more than one occasion, usually at 1-year intervals. Blood was collected before breast examination, between 9:00 a.m. and 3:00 p.m., and at any time during the menstrual cycle of the premenopausal women. No fasting was required. Blood specimens were kept at room temperature for approximately 1 h and at 4°C for 30 min. Samples were then centrifuged at 3500 rpm for 15 min, and serum was partitioned into 1-ml aliquots in airtight plastic vials and frozen at -80°C for long-term storage.

Study Groups. Since prolactin levels drop after the menopause, separate groups of premenopausal and postmenopausal women were studied. Three groups were selected: (a) premenopausal women; (b) postmenopausal women who gave repeat samples in any phase of the menstrual cycle; and (c) postmenopausal women who gave multiple samples in the luteal phase. The third group was chosen to control for the additional variability in premenopausal prolactin levels caused by menstrual cycling. A sample was considered to be luteal phase if it was drawn within 11 days of the next menstruation. Date of next menstruation was obtained from mailbox calendars distributed at the time of blood drawing. Statistical analyses were carried out separately in each of the three study groups.

To control for variability of prolactin levels with age (5), age at sampling was restricted to ranges within which prolactin levels are relatively constant. Thus, premenopausal samples had to have been drawn between the ages of 55 and

Received 11/18/92; revised 2/10/93; accepted 2/10/93.

1 This research was supported by Grants CA 13341, CA 16087 and CA 34588 from the National Cancer Institute.
2 To whom requests for reprints should be addressed, Institute of Environmental Medicine, New York University Medical Center, 341 East 25th Street, New York, NY 10010.
3 The abbreviations used are: NYU, New York University; p, intraclass correlation coefficient.
Reliability of Serum Prolactin Measurements

69 years, and premenopausal samples had to have been drawn between the ages of 34 and 49 years. Only women with at least one intact ovary were eligible. Any samples drawn in the 6 months following pregnancy were excluded. Also excluded were samples drawn while the subject was taking drugs that might affect serum levels of prolactin. These drugs included sex hormones; glucocorticoids; major tranquilizers, such as haldol and prolixin; certain blood pressure medications (reserpine, α-methyldopa, pargylne); certain antidepressants, such as tofranil and pamelor; codeine; apomorphine; dilantin; l-dopa; metoclopramide; amphetamines; dopamine; bromocryptine; and cimetidine (4, 6-8).

Laboratory Assays. The serum samples were packed in dry ice and sent to the Netherlands Cancer Institute, where they were stored at −80°C until they were assayed in a single batch. The laboratory assay was an automated "sandwich" enzyme immunoassay technique using two monoclonal antibodies highly specific to prolactin (9). Boehringer Mannheim (Germany), the manufacturer, reports no measurable cross-reaction with other anterior pituitary hormones (hGH, hCG, TSH, FSH, and LH). All sample values were above the 2 ng/ml lower limit of detection of the assay.

Reliability. Our goal was to determine how well a single serum measurement characterizes a woman’s long-term serum level of prolactin relative to other women. We assume that each woman has a long-term mean level of prolactin, about which her individual measurements fluctuate. Part of this fluctuation is due to error in the laboratory measurement and part to real biological variation in prolactin levels over time. We assumed the following model (model 1) for \( Y_{jk} \), the measured prolactin value for the \( k \)th aliquot from the \( j \)th blood sample donated by the \( j \)th subject:

\[
Y_{jk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}
\]

where \( \mu \) is the overall mean prolactin level, \( \alpha_i \) is the mean prolactin level for the \( i \)th subject, \( \beta_j \) is the difference between the true prolactin level for subject \( i \) at donation \( j \) and her mean level, and \( \epsilon_{ijk} \) is the laboratory measurement error, or the difference between the measured and true prolactin values for the \( k \)th aliquot from the \( j \)th sample donated by subject \( i \). The random variables \( \alpha_i, \beta_j, \epsilon_{ijk} \) were assumed to be mutually independent with mean 0 and variance \( \sigma^2_\alpha, \sigma^2_\beta, \sigma^2_\epsilon \) respectively. Under these conditions, the variance of a prolactin measurement (\( Y_{ijk} \)) is the sum of the three variance components, \( \sigma^2_\alpha, \sigma^2_\beta, \sigma^2_\epsilon \). The total variance of the mean levels of the subjects (the between-subjects component), while \( \sigma^2_\epsilon \) represents the variance of the distribution of the mean levels of the subjects (the between-subjects component), while \( \sigma^2_\epsilon \) is the error component of variance due to laboratory error (\( \sigma^2_\epsilon \)). The only analysis in which the duplicate measurements were used was the analysis of variance model (2)

\[
Y_{ij} = \mu + \alpha_i + \delta X_j + \epsilon_{ij}
\]

where \( \alpha_i \) is the random subject effect, \( \delta \) is the within-subject error (analogous to the sum of \( \beta_j \) and \( \epsilon_{ijk} \) in model 1 but adjusted for the covariates \( X_j \)), and \( \epsilon_{ij} \) is the log of the measured prolactin value for subject \( i \) at donation \( j \). In this model the fixed effect \( \delta \) assumed constant across subjects, measures the change in log prolactin value for a 1-unit increase in the value of the covariate. The effect of categorical covariates was assessed using the analysis of variance model.
(model 3)

\[ Y_{ijk} = \mu + \alpha_i + \delta_j + \omega_{ijk} \]

where \( \alpha_i \) is the random subject effect, \( \omega_{ijk} \) is the within-subject error, \( \delta_j \) is the (fixed) effect on the log prolactin value of level \( j \) of the covariate, and \( Y_{ijk} \) is the measured log prolactin value for the \( k \)th sample donated by subject \( i \) at level \( j \) of the covariate. The categorical covariates investigated were time of day of the sample (before versus after noon), time since last meal (<3 versus ≥3 h), and phase of the menstrual cycle. Phase of cycle, analyzed in the group of premenopausal subjects with samples taken throughout the menstrual cycle, was categorized as luteal (0–11 days prior to the next menstruation), ovulatory (12–16 days), late follicular (17–19 days), and early follicular (≥20 days). The covariate effects (β) for models 2 and 3 were estimated using the Statistical Analysis System (SAS) General Linear Models procedure (14). Variance components were estimated using the SAS VARCOMP procedure, and reliability coefficients adjusted for covariates were computed as (see Refs. 10 and 15)

\[ \frac{\sigma^2_{\alpha} + \sigma^2_{\omega}}{\sigma^2_{\alpha}} \]

Results

Serum prolactin values in the three study groups are given in Table 1. The prolactin values in the postmenopausal women ranged from 2.9 to 43.2 ng/ml with a geometric mean of 7.8. The premenopausal prolactin values were higher, ranging from 3.5 to 80.4 ng/ml with a geometric mean of 12.1. The geometric means in all groups were less than the arithmetic means, indicating that the distributions were skewed to the right. Values for the premenopausal group are also shown with one influential subject (number 21) excluded. Four of the five highest measurements in the postmenopausal group belonged to this subject.

Table 1  Serum prolactin values in the three study groups

<table>
<thead>
<tr>
<th></th>
<th>Postmenopausal</th>
<th>Premenopausal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All subjects</td>
<td>Minus subject 21</td>
</tr>
<tr>
<td>Age at sampling (years)</td>
<td>55-69</td>
<td>34-49</td>
</tr>
<tr>
<td>Number of subjects</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Number of replicates</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Prolactin values (ng/ml)</td>
<td>2.9-4.2</td>
<td>2.9-32.5</td>
</tr>
<tr>
<td>Arithmetic mean</td>
<td>8.9</td>
<td>8.2</td>
</tr>
<tr>
<td>Geometric mean</td>
<td>7.8</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Variance components

Between subjects 0.166 0.107 0.101 0.105
Within subjects 0.052 0.052 0.110 0.116
Within samples (laboratory) 0.011

Intraclass correlation

95% confidence interval

0.66-0.85
0.31-0.62

Coefficient of variation

(laboratory) 4.8%

As expected since the age ranges were chosen to minimize the effect of age. Sample storage time, which ranged from 0.5 to 5.8 years, likewise did not affect the prolactin value: the direction of the storage time effect was not consistent in the three study groups nor was the effect statistically significant in any group, implying that prolactin did not undergo degradation during storage. Time of day and time since last meal influenced prolactin levels only in the postmenopausal group. With both time of day and time since last meal in the model (i.e., with each variable controlled for the other), afternoon prolactin values in postmenopausal women were estimated to be 17% higher than morning values (\( P = 0.001 \)), and drawing the sample within 3 h of eating was estimated to raise the prolactin value 10.5% (\( P = 0.04 \)). In spite of the significance of these effects, adjustment for both time of day and time since last meal increased the log-scale reliability coefficient in postmenopausal women only marginally, from 0.76 to 0.78. It is likely that time of day would have been more influential if evening or night samples had been included.
Discussion
These results indicate that the ability of a single measurement to characterize a woman’s long-term serum prolactin level relative to other women depends on her menopausal status. For the postmenopausal subjects the reliability of a single log prolactin measurement was quite high (0.76), which implies that prolactin levels are quite stable after menopause. For the premenopausal subjects the reliability of a single log measurement was considerably worse (0.48), and negligible improvement was observed after controlling for phase of the menstrual cycle. Because prolactin levels are more variable in premenopausal women, several samples are needed to raise the reliability of the geometric mean to the level observed in postmenopausal women. Three premenopausal samples are needed to raise the reliability to 0.73 (10), four samples to raise the reliability to 0.79.

Variability due to the laboratory assay was reduced in this study by processing all samples in a single batch. If different samples were assayed at different times and if significant random batch-to-batch variation were present, the reliability of the measurements would be somewhat lower than the values calculated here. However, since we have observed systematic time trends in the assay results for several hormones under investigation in our case-control study, we believe that batch effects are likely to vary systematically over time and between laboratories. By assaying all samples in the same batch, we avoided any systematic time trends in the assay results. We are employing a similar approach in our case-control study by using a matched design and assaying all samples from a case and her matched controls in the same batch.

It should be noted that this study contained few premenopausal women. Our requirement that phase of cycle be known for all premenopausal measurements excluded most premenopausal samples, which were likely to have phase of cycle unknown. Since control for phase of cycle had little influence on the premenopausal reliability estimates, we would expect the reliability of a prolactin measurement in menstruating premenopausal women to be similar to the premenopausal estimate. We had no direct information on the reliability of prolactin measurements taken shortly after menopause, since only six measurements in the postmenopausal group were taken within 4 years of last period. However, it seems reasonable to assume that the reliability of a prolactin measurement taken shortly after menopause lies between the premenopausal and postmenopausal estimates.

We were surprised that the reliability of the premenopausal prolactin measurements was not improved when phase of the menstrual cycle was controlled. Analyses of the premenopausal, any-phase group demonstrated that geometric mean prolactin levels followed a predictable pattern through the cycle. But these cyclic changes were small compared with the range of the measurements. It is only during the brief ovulatory phase that the mean prolactin level rises appreciably. Most measurements were therefore taken during a time in the cycle when the mean level was relatively constant. The lack of reduction in the within-subjects variability in the luteal phase group compared with the unrestricted group suggests that changes in prolactin levels across the menstrual cycle are small compared with the variability in the prolactin level of an individual subject from one luteal phase to another.

The consequences of unreliability are a loss of statistical power and a bias toward unity in relative risk estimates. The effective sample size is reduced by a factor of \( r^2 \) and the log odds ratio in the logistic regression model is attenuated by this same factor (17). The reliability of 0.76 for single postmenopausal prolactin measurements is close to the value of 0.80 which is considered acceptable for relative risk estimation in epidemiological studies (18). We therefore conclude that a single serum sample is sufficiently characteristic of the mean prolactin level of a postmenopausal woman to provide meaningful exposure data for epidemiological research. For premenopausal women, however, four samples are required to raise the reliability of the geometric mean to a target level near 0.80. Since controlling for phase of the menstrual cycle produced only slight improvement in the reliability of the premenopausal prolactin measurements, it appears that control for phase of cycle is unnecessary in epidemiological analyses of serum prolactin in premenopausal women.

Acknowledgments
The authors thank Dr. Mimi Kim for statistical advice, Miriam Moseben for helpful information, and Frances Mastrosa, Anna Bassiri, Dorothe Linders, Loes de Groot-Beckers, and Tiny Korse for technical assistance. The assistance of the Gutman Breast Diagnostic Institute in all phases of data collection and of its Medical Director Emeritus, Dr. Philip Strax, is also gratefully acknowledged.

References
Reliability of serum prolactin measurements in women.
K L Koenig, P Toniolo, P F Bruning, et al.


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/2/5/411

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.