Research

Reliability of follicle-stimulating hormone measurements in serum

Alan A Arslan*[1], Anne Zeleniuch-Jacquotte[1], Annekatrin Lukanova[2], Sabina Rinaldi[2], Rudolf Kaaks[2] and Paolo Toniolo[1]

Address: [1]New York University School of Medicine, New York, NY, USA and [2]Hormones and Cancer Group, International Agency for Research on Cancer, Lyon, France

Email: Alan A Arslan* - akhmea01@med.nyu.edu; Anne Zeleniuch-Jacquotte - jacqua02@med.nyu.edu; Annekatrin Lukanova - lukanova@iarc.fr; Sabina Rinaldi - rinaldi@iarc.fr; Rudolf Kaaks - kaaks@iarc.fr; Paolo Toniolo - toniop01@gcrc.med.nyu.edu

* Corresponding author

Abstract

Background: Follicle-stimulating hormone (FSH), a member of gonadotropin family, is critical for follicular maturation and ovarian steroidogenesis. Serum FSH levels are known to fluctuate during different phases of menstrual cycle in premenopausal women, and increase considerably after the menopause as a result of ovarian function cessation. There is little existing evidence to guide researchers in estimating the reliability of serum FSH measurements. The objective of this study was to assess the reliability of FSH measurement using stored sera from an ongoing prospective cohort – the NYU Women's Health Study.

Methods: Sixty healthy women (16 premenopausal, 44 postmenopausal), who donated at least two blood samples at approximately 1-year intervals were studied. An immunoradiometric assay using a sandwich monoclonal antibodies technique was used to measure FSH levels in serum.

Results: The reliability of a single log-transformed FSH measurement, as determined by the intraclass correlation coefficient, was 0.70 for postmenopausal women (95% confidence interval (CI), 0.55–0.82) and 0.09 for premenopausal women (95% CI, 0–0.54).

Conclusions: These results suggest that a single measurement is sufficient to characterize the serum FSH level in postmenopausal women and could be a useful tool in epidemiological research. For premenopausal women, however, the reliability coefficient was low, suggesting that a single determination is insufficient to reliably estimate a woman's true average serum FSH level and repeated measurements are desirable.

Background

Follicle-stimulating hormone (FSH) plays a key role in the development and function of the reproductive system and is widely used both in clinical and research settings. The accurate and reliable measurement of FSH levels is essential for safe and successful treatment in developmental and reproductive medicine [1], as well as for research studies examining the association between FSH levels and various disease outcomes.

FSH is a member of the gonadotropin family, which includes also luteinizing hormone (LH) and human chorionic gonadotropin (hCG). Gonadotropins are complex heterodimeric glycoproteins which consist of two linked protein components designated as the α- and β-subunits.

The α-subunit is common to the three gonadotropins, whereas the β-subunit confers specificity and biological activity.

According to the "two cell, two gonadotropin" theory [2–5], both FSH and LH are necessary for ovarian follicular maturation and the syntheses of ovarian steroid hormones. LH promotes the production of androgens (dehydroepiandrosterone, androstenedione, and testosterone) from cholesterol and pregnenolone, by stimulating 17α-hydroxylase activity in the thecal cells. The androgens then diffuse to the granulosa cells where FSH stimulates the expression of the cytochrome P450 aromatase, which converts the androgens to estrogens [6,7].

The measurement of FSH in circulation is employed in the diagnosis of disorders of reproduction and development, whereas therapeutic preparations of FSH are widely used for induction of ovulation in women and stimulation of spermatogenesis in men [1]. The effects of gonadotropins may not be limited to endocrine and reproductive functions. Excessive gonadotropin stimulation of the ovarian epithelium has been postulated to be one of the possible mechanisms of ovarian carcinogenesis [8]. However, studies that directly examined the association between serum levels of gonadotropins and ovarian cancer risk have not been consistent with this theory [9,10].

Before starting complex epidemiological studies examining the associations between FSH and various diseases, it is important to assess the extent of the hormone's underlying fluctuations in circulation. FSH levels peak during the menstruation and ovulatory phase and are lower during the late follicular and luteal phases of the menstrual cycle. After menopause, FSH levels gradually increase through negative biofeedback as a result of ovarian function cessation. Given a substantial fluctuation of FSH levels under normal physiological conditions, determination of FSH in a single measurement may provide inadequate estimates of the true average values over extended periods of time.

The purpose of this preliminary study was to examine the reliability of FSH measurements in premenopausal and postmenopausal women using a subset of subjects with repeated serum FSH measurements from the New York University (NYU) Women's Health Study.

Methods
Study Population
Between March 1985 and June 1991, the NYU Women's Health Study enrolled a cohort of 14,275 women aged 34–65 years, attending a breast screening clinic in New York City [11,12]. The cohort was restricted to women who in the preceding 6 months were neither pregnant nor treated with hormones. At the time of enrollment and at annual screening visits thereafter, subjects were asked to complete questionnaires on medical, anthropometric, reproductive, and dietary factors and to provide 30 mL of non-fasting peripheral venous blood. Approximately half of the participants gave blood at repeated visits, on average at 1-year intervals. After blood drawing, tubes were kept covered at room temperature (20°C) for 15 minutes, then at 4°C for 60 minutes to allow clot retraction, and then centrifuged at 3500 rpm for 15 minutes. Supernatant serum was partitioned into 1-mL aliquots and immediately stored at -80°C for future biochemical analyses.

The subjects for this study were selected from a subset of the NYU Women's Health Study subjects participating in a multi-center nested case-control study of body mass index in relation to ovarian cancer [13] for whom serum FSH levels were measured to verify their self-reported menopausal status. Women were classified as premenopausal if they reported at least one menstrual cycle during the 6 months prior to enrollment. The number of days prior to the next menses and the phase of the cycle were calculated using calendars that the study subjects were instructed to mark and return following their next menses after enrollment. Women were classified as postmenopausal if they reported absence of menstrual cycles in the previous 6 months, a total bilateral oophorectomy, or a hysterectomy without total oophorectomy if their age was 52 years or older. Postmenopausal status was confirmed by serum FSH level greater than 12.5 mIU/mL, as previously described [13]. The subjects eligible for the current study included women, who were free of cancer and who had repeated FSH measurements within at least one-year interval. A total of sixty healthy women: 16 premenopausal (secretory phase), 44 postmenopausal, who have met these eligibility criteria were selected to study the reliability of FSH measurements in serum.

Laboratory Methods
Biochemical analyses of serum samples were performed at the Laboratory of Hormones and Cancer Group, International Agency for Research on Cancer (IARC, Lyon, France). Laboratory personnel were unaware of subjects' case-control status and of the temporal sequence of the samples. Repeated samples from the same subjects were always analyzed in the same batch. Serum FSH was measured by an immunoradiometric assay (FSH IRMA, Diagnostic System Laboratories, TX, USA). The FSH IRMA is a non-competitive assay in which the analyte is sandwiched between two monoclonal antibodies. The first antibody is coated on the walls of the tubes used in the analysis, while the second antibody is radiolabeled for detection. The unbound fraction is removed by a washing step. The amount of radioactivity counted in the assay tubes is directly proportional to the amount of analyte in the

Table 1: FSH levels and age at two visits in subset of healthy women, NYU Women's Health Study

| | Premenopausal *N* = 16 | | Postmenopausal *N* = 44 | |
	Visit 1	Visit 2	Visit 1	Visit 2
FSH values, mIU/mL				
Median	2.6	4.5	43.1	44.1
Range	0.4–8.9	1.2–19.1	13.5–71.1	12.9–77.1
Mean	3.4	5.7	43.3	43.9
Standard deviation	2.6	4.2	13.6	14.4
Age at sampling, years				
Median	42	43	58	59
Range	36–56	37–57	43–65	45–66
Mean	43.7	45.0	57.4	59.0
Standard deviation	5.0	4.8	4.7	4.8

sample. A set of standards with known amount of FSH is used to plot a standard curve from which the amount of FSH in the samples can be calculated. *Assay sensitivity:* all reported sample values were above the 0.11 mIU/mL lower detection limit. *Assay specificity:* the FSH kit manufacturer (DSL, Texas, USA) reported no measurable cross-reaction with other gonadotropins (LH and hCG). The within-assay coefficients of variation provided by the IARC laboratory were ranging from 3.2 to 4.6% depending on the serum FSH concentration.

Statistical Methods

The reliability was assessed by calculating the intraclass correlation coefficients (reliability coefficients). Variance components were estimated in an ANOVA analysis assuming a one-way random effects model [14]. Computations were performed on the natural log-transformed data in order to reduce the positive skewness of the raw data. Exact 95% confidence intervals (CIs) were calculated as described by Shrout and Fleiss [15]. A mixed effects regression model was used to assess whether age, body mass index (BMI = weight in kg/height in m^2) or storage time were predictive of FSH serum level. Spearman correlation coefficients between FSH levels and these three explanatory variables were also computed, separately for each visit.

Results

Of the sixty women with repeated FSH measurements included in the study, 16 were premenopausal and 44 were postmenopausal. Premenopausal women had a mean age (± SD) at first blood donation of 43.7 (± 5.0) years and a mean BMI of 25.5 kg/m^2 (± 3.0 kg/m^2). For postmenopausal women, the mean age was 57.4 years (± 4.7 years) and the mean BMI was 25.9 kg/m^2 (± 4.3 kg/m^2). Mean times in storage of the serum samples were 14.0 years (± 0.5 years) and 12.6 years (± 1.0 years) for the first and subsequent visits, respectively.

Serum FSH values of the study subjects at enrollment and during a repeat visit are presented in Table 1, according to menopausal status. The FSH levels in the premenopausal women at enrollment ranged from 0.4 to 8.9 mIU/mL with a median of 2.6 mIU/mL. During the repeat visit approximately one year later, their serum FSH levels were higher, ranging from 1.2 to 19.1 mIU/mL with a median of 4.5 mIU/mL. As expected, the baseline postmenopausal FSH values were substantially higher, as compared to premenopausal women, ranging from 13.5 to 71.1 mIU/mL with a median of 43.1 mIU/mL.

Variance components and the resulting intraclass correlation (reliability) coefficients computed on the natural log-transformed data are shown in Table 2. Among premenopausal women, the reliability coefficient for FSH was 0.09 (95% CI 0.0–0.54). The reliability coefficient for postmenopausal women was 0.70 (95% CI 0.55–0.82). The lower reliability in premenopausal women appears to be related to greater within-subject variability before menopause, since the within-subject variance component was 15-fold higher in the premenopausal women compared to the postmenopausal group. In addition, the between-subject variance component was greater in the postmenopausal group than in the premenopausal women, suggesting that this variance component may also contribute to the observed differences in serum FSH reliability before and after menopause.

Using a mixed effects regression model, neither age nor storage duration were predictive of FSH level. There was a marginally significant negative association between BMI and FSH level (p = 0.045). The Spearman correlation coefficient between FSH level and BMI was -0.12 at the first visit and -0.27 at the second visit.

Variability in the results of the laboratory assay (the within-assay coefficient of variation) was only a small

Table 2: Between- and within-subject variance and reliability coefficients for serum FSH in subset of healthy women, NYU Women's Health Study

	Premenopausal *N* = 16	Postmenopausal *N* = 44
Variance components		
Between subjects	0.057	0.102
Within subjects	0.600	0.044
Reliability coefficient, ρ	0.09	0.70
95% confidence interval	0–0.54	0.55–0.82
Coefficient of variation, % (within-assay)	4.6	3.2

proportion of the total variability. The within-assay coefficient of variation provided by the laboratory was ranging from 3.2 to 4.6% depending on the serum FSH concentration (Table 2).

Discussion

The results demonstrate that the ability of a single measurement to characterize a woman's long-term serum FSH level depends on her menopausal status. For the postmenopausal women, the reliability of a single FSH measurement was adequate (reliability coefficient = 0.70, 95% CI 0.55–0.82), which suggests that serum FSH levels are fairly stable after menopause. For the premenopausal women, the reliability of a single FSH measurement was considerably worse (reliability coefficient = 0.09, 95% CI 0–0.54), which implies that FSH levels are more variable in premenopausal women.

Variability due to the laboratory assay was reduced in the current study by processing the samples from the same subjects in the same batch. The low within-assay coefficients of variation (<5%) suggest that most of the within-subject variability was the expression of real biological variation and not due to laboratory measurement error. If the samples were assayed at different times and if substantial random batch-to-batch variation were present, the reliability of the measurement would be somewhat lower than the values presented here.

Variation in FSH levels during the menstrual cycle is thought to be critical in the mechanism of FSH-dependent selection of the dominant follicles [16] and could affect the reliability estimates in premenopausal women. Lack of storage time effect suggests that degradation of specimens during long-term freezer-storage seems an unlikely explanation for low reliability of FSH in premenopausal women in the current study. In addition, the higher mean FSH levels during the repeated visits and the mean age of premenopausal women during the repeated visit (45 years) suggest that some of these women may have been perimenopausal.

Although the ovarian function markedly decreases after the menopause, gonadotropins may play a role in postmenopausal women. Ovarian tissues from postmenopausal women express gonadotropin receptors [17] and could synthesize steroid hormones [18,19]. Longcope [20] has shown that postmenopausal ovary is characterized by a markedly decreased secretion of estrogens and certain androgens, but the secretion of testosterone is preserved to a large extent in most postmenopausal women. These observations suggest that gonadotropin response and endocrine function do not entirely cease after the menopause and may play a role in certain hormone-dependent conditions, such as polycystic ovaries [21].

The importance of assessing the reliability of exposure measurement prior to planning the complex epidemiological investigations is based on the fact that poor reliability may reduce the effective sample size [22], result in a loss of statistical power and a bias toward unity in relative risk estimates [23]. The issue of reliability is even more important for cohort studies utilizing prospectively collected biological samples, where strategies to preserve the valuable specimens for only reliable exposure measurements should be given consideration.

In conclusion, the results of this study indicate that whereas a single determination could be inadequate to reliably estimate serum FSH in premenopausal women, a single measurement may be sufficiently characteristic of the average serum FSH levels in postmenopausal women. These results have implications for the design of epidemiological studies to evaluate the role of FSH in various hormone-dependent conditions.

List of abbreviations

FSH, follicle-stimulating hormone; LH, luteinizing hormone; hCG, human chorionic gonadotropin; NYU, New York University; IRMA, immunoradiometric assay; mIU/mL, milli-international units per milliliter; ANOVA, analysis of variance; BMI, body mass index; CI, confidence interval; SD, standard deviation.

Competing interests

None declared.

Authors contributions

Acknowledgments

References

1. Rose MP, Gaines Das RE and Balen AH: **Definition and measurement of follicle stimulating hormone** *Endocr Rev* 2000, **21:**5-22.
2. Fevold HL: **Synergism of follicle stimulating and luteinizing hormones in producing estrogen secretion** *Endocrinology* 1941, **28:**33-36.
3. Greep RO, Van Dyke HB and Chow BF: **Gonadotropin of swine pituitary: various biological effects of purified thylkentrin (FSH) and pure matakentrin (ICSH)** *Endocrinology* 1942, **30:**635-649.
4. Kobayashi M, Nakano R and Ooshima A: **Immunohistochemical localization of pituitary gonadotrophins and gonadal steroids confirms the 'two-cell, two-gonadotrophin' hypothesis of steroidogenesis in the human ovary** *J Endocrinol* 1990, **126:**483-488.
5. Hillier SG, Whitelaw PF and Smyth CD: **Follicular oestrogen synthesis: the 'two-cell, two-gonadotrophin' model revisited** *Mol Cell Endocrinol* 1994, **100:**51-54.
6. Erickson GF, Magoffin DA, Dyer CA and Hofeditz C: **The ovarian androgen producing cells: a review of structure/function relationships** *Endocr Rev* 1985, **6:**371-399.
7. Richards JS: **Hormonal control of gene expression in the ovary** *Endocr Rev* 1994, **15:**725-751.
8. Cramer DW and Welch WR: **Determinants of ovarian cancer risk. II. Inferences regarding pathogenesis** *J Natl Cancer Inst* 1983, **71:**717-721.
9. Helzlsouer KJ, Alberg AJ, Gordon GB, Longcope C, Bush TL, Hoffman SC and Comstock GW: **Serum gonadotropins and steroid hormones and the development of ovarian cancer** *JAMA* 1995, **274:**1926-1930.
10. Akhmedkhanov A, Toniolo P, Zeleniuch-Jacquotte A, Pettersson KS and Huhtaniemi IT: **Luteinizing hormone, its beta-subunit variant, and epithelial ovarian cancer: the gonadotropin hypothesis revisited** *Am J Epidemiol* 2001, **154:**43-49.
11. Toniolo PG, Pasternack BS, Shore RE, Sonnenschein E, Koenig KL, Rosenberg C, Strax P and Strax S: **Endogenous hormones and breast cancer: a prospective cohort study** *Breast Cancer Res Treat* 1991, **18 Suppl 1:**S23-S26.
12. Toniolo PG, Levitz M, Zeleniuch-Jacquotte A, Banerjee S, Koenig KL, Shore RE, Strax P and Pasternack BS: **A prospective study of endogenous estrogens and breast cancer in postmenopausal women** *J Natl Cancer Inst* 1995, **87:**190-197.
13. Lukanova A, Toniolo P, Lundin E, Micheli A, Akhmedkhanov A, Muti P, Zeleniuch-Jacquotte A, Biessy C, Lenner P, Krogh V, Berrino F, Hallmans G, Riboli E and Kaaks R: **Body mass index in relation to ovarian cancer: a multi-centre nested case-control study** *Int J Cancer* 2002, **99:**603-608.
14. Donner A: **Inference for the intraclass correlation coefficients** *Int Stat Rev* 1986, **54:**67-82.
15. Shrout PE and Fleiss JL: **Intraclass correlations - uses in assessing rater reliability** *Psychol Bull* 1979, **86:**420-428.
16. Scheele F and Schoemaker J: **The role of follicle-stimulating hormone in the selection of follicles in human ovaries: a survey of the literature and a proposed model** *Gynecol Endocrinol* 1996, **10:**55-66.
17. Kobayashi M, Nakano R and Shima K: **Immunohistochemical localization of pituitary gonadotropins and estrogen in human postmenopausal ovaries** *Acta Obstet Gynecol Scand* 1993, **72:**76-80.
18. Dennefors BL, Janson PO, Knutson F and Hamberger L: **Steroid production and responsiveness to gonadotropin in isolated stromal tissue of human postmenopausal ovaries** *Am J Obstet Gynecol* 1980, **136:**997-1002.
19. Mattingly RF and Huang WY: **Steroidogenesis of the menopausal and postmenopausal ovary** *Am J Obstet Gynecol* 1969, **103:**679-693.
20. Longcope C: **Endocrine function of the postmenopausal ovary** *J Soc Gynecol Investig* 2001, **8:**S67-S68.
21. Mechain C, Cedrin I, Pandian C and Lemay A: **Serum FSH bioactivity and response to acute gonadotrophin releasing hormone (GnRH) agonist stimulation in patients with polycystic ovary syndrome (PCOS) as compared to control groups** *Clin Endocrinol (Oxf)* 1993, **38:**311-320.
22. McKeown-Eyssen GE and Tibshirani R: **Implications of measurement error for the sample sizes of case-control studies** *Am J Epidemiol* 1994, **139:**415-421.
23. de Klerk NH, English DR and Armstrong BK: **A review of the effects of random measurement error on relative risk estimates in epidemiological studies** *Int J Epidemiol* 1989, **18:**705-712.