Borrower: WSM

Lending String: *ME6,CHS,ONE,MXC,MCL
Patron: Glaser, Rebecca-122041
Journal Title: Cancer epidemiology, biomarkers & prevention ; a publication of the American Association for Cancer Research /
ISSN: 1055-9965
System: OCLC
PMID: 23441645

Volume: 16 **Issue:** 9
Month/Year: Sept 2007 **Pages:** 1713-19

Article Author:

Article Title: Stanczyk; Standardization of Steroid hormone assays; Why, How, and When?

Imprint: Philadelphia, PA ; American Association

ILL Number: 33819072

Location:

Odyssey: 206.107.42.197
ARIEL No
Ariel:
Regular
Charge
Maxcost: 15.00IFM

Shipping Address:
Fordham Health Sciences Library
Wright State University- ILL
3640 Colonel Glenn Highway
Dayton, OH 45435-0001
Billing Address:
Fordham Health Sciences Library
Fax:
Email:
Phone:
Notes: Borrowing Notes;

Commentary

Standardization of Steroid Hormone Assays: Why, How, and When?

Frank Z. Stanczyk,[1] Jennifer S. Lee,[2] and Richard J. Santen[3]

[1]Departments of Obstetrics and Gynecology, and Preventive Medicine, University of Southern California, Keck School of Medicine, Women's and Children's Hospital, Los Angeles, California; [2]Division of Endocrinology, Clinical Nutrition, and Vascular Medicine, Department of Internal Medicine, University of California at Davis, Sacramento, California; and [3]Clinical Research, Cancer Center, University of Virginia, Charlottesville, Virginia

Abstract

Lack of standardization of high-quality steroid hormone assays is a major deficiency in epidemiologic studies. In postmenopausal women, reported levels of serum 17β-estradiol (E_2) are highly variable and median normal values differ by approximately a 6-fold factor. A particular problem is the use of E_2 assays for prediction of breast cancer risk and osteoporotic fractures, where assay sensitivity may be the most important factor. Identification of women in the lowest categories of E_2 levels will likely provide prognostic information that would not be available in a large group of women in whom E_2 levels are undetectable by less sensitive assays. Detailed and costly methods involving extraction and chromatography in conjunction with RIA provide generally acceptable E_2 results in postmenopausal serum, whereas less tedious, direct immunoassays suffer from inadequate specificity and sensitivity. Studies comparing the two types of methods generally report higher E_2 values with the direct methods as a result of cross-reactivity with other steroids and reduced correlation with biological variables such as body mass index. Similar problems exist with measurements of E_2 and estrone in men, and estrone and testosterone in women. Interest in mass spectrometry–based assays is increasing as potential gold standard methods with enhanced sensitivity and specificity; however, these assays require costly instrumentation and highly trained personnel. Taking all of these issues into consideration, we propose establishment of standard pools of premenopausal, postmenopausal, and male serum, and utilization of these for cross-comparison of various methods on an international basis. An oversight group could then establish standards based on these comparisons and set agreed upon confidence limits of various hormones in the pools. These criteria would allow validation of sensitivity, specificity, precision, and accuracy of current steroid hormone assay methodology and provide surrogates until a true gold standard can be developed.(Cancer Epidemiol Biomarkers Prev 2007;16(9):1713–9)

Introduction

High-quality steroid hormone assays with very good sensitivity, specificity, and reproducibility are essential to the validity of epidemiologic studies. Steroid hormone measurements play a critical role in a variety of studies. Among these are epidemiologic investigations of many major diseases, including osteoporotic fracture, cognitive dysfunction, and hormonally related cancers, including cancers of the breast, ovary, endometrium, prostate, and testes. However, epidemiologic studies use many different assay methods with varying performance. This is evident, for example, in a survey of recent large epidemiologic studies of bone health in postmenopausal women (Table 1; refs. 1-6). Use of different assay methods with varying performance contributes problematically to inconsistent results and varied interpretations across epidemiologic studies.

The lack of standardization of high-quality steroid hormone assays is a major deficiency in epidemiologic studies, resulting in varying findings in hormone concentrations and hindering the ability to draw definitive quantitative conclusions. This problem is particularly relevant regarding low levels of steroid hormones such as 17β-estradiol (E_2) and estrone (E_1) in postmenopausal women and men, and testosterone in women. The majority of E_2 assays suffer from lack of sufficient sensitivity, specificity, precision, and/or accuracy. No gold standard exists to allow objective validation and cross comparisons among various assays to ensure maximal quality control. An example in which this problem is particularly significant is in assessing breast cancer risk associated with low E_2 levels (<30 pg/mL) in postmenopausal women. To clarify the relation between endogenous E_2 levels and breast cancer risk in the setting of varied assay methods, limited statistical power, modest effect estimates, and limited variation in hormone levels across individual studies, a pooled analysis of nine epidemiologic studies of endogenous

Received 9/10/06; revised 6/19/07; accepted 7/2/07.

Requests for reprints: Frank Z. Stanczyk, Departments of Obstetrics and Gynecology, and Preventive Medicine, University of Southern California, Keck School of Medicine, Women's and Children's Hospital, Room 1M2, 1240 North Mission Road, Los Angeles, CA 90033. Phone: 323-226-3220; Fax: 323-226-2850; E-mail: fstanczyk@socal.rr.com

doi:10.1158/1055-9965.EPI-06-0765

postmenopausal hormones and breast cancer risk was conducted (7). This study highlighted both the approximate 6-fold difference in median values among studies (range of median values in normal postmenopausal women, 6.0-37 pg/mL), and the wide range of detection limits (from 0.8 to 10 pg/mL) for the E_2 assay methods used. Similar issues exist for measurement of serum levels of E_2 and E_1 in studies of men with benign prostatic hyperplasia and prostate cancer.

These discrepancies in assay performance particularly limit investigations where comparisons of absolute, rather than relative, values of sex hormone concentrations are needed. Currently, estimating disease risk in relation to serum sex steroids relies on comparing relative risks of disease across tertiles or quartiles of sex steroid concentrations. This epidemiologic approach relies less on an assay producing accurate measurements of absolute concentration values. A sex steroid measurement will likely be classified into the correct tertile or quartile category as long as it is not near the border of two categories, thereby helping to circumvent assay inaccuracies. However, direct comparisons of sex hormone data across epidemiologic studies and recommendations about the use of hormone assays in the clinical setting would be facilitated by the use of assays across studies that provide both accurate and precise measurements. Determining accurate absolute concentration values becomes essential, for example, to examine dose-response relationships with exogenous hormone use or effects of clinically relevant hormonal threshold levels for disease outcomes. Current assay limitations potentially hinder further understanding of the biological significance of serum sex steroids and the ability to make clinical recommendations about hormonal strategies for prevention and treatment, based on actual concentration values. Furthermore, misleading information may be obtained if no attempt is made to determine how well assay measurements correlate against those obtained by a gold standard assay. This is especially true when commercial kits are used for measuring steroid hormones in epidemiologic studies. Assays done with such kits are rarely validated thoroughly by the kit manufacturer.

The purpose of this commentary is to (*a*) point out the current limitations of steroid hormone assays using serum or plasma, particularly in measurements of low concentrations of endogenous hormones (e.g., E_2 or testosterone in women and E_2 in men); (*b*) provide a perspective as to potential approaches to assay standardization; and (*c*) project when standardization of such assays may be accomplished.

Immunoassay Methods

Extraction/Chromatographic RIAs. Most of our knowledge about the physiologic and diagnostic role of steroid hormones in women and men is derived from studies in which circulating levels of these compounds were measured by radioimmunoassay (RIA) methods. The basis for the first RIA method was described in 1959 by Yalow and Berson, who showed that [^{131}I]insulin could be displaced by nonradioactive insulin from insulin-binding protein, and that [^{131}I]insulin was inversely and quantitatively related to the total amount of insulin present (8). Subsequently, specific antisera to human insulin were prepared and resulted in the first RIA, with sufficient sensitivity to detect endogenous insulin in human blood (9). A new era in reproductive endocrinology was launched when Odell et al., in 1967, developed the first RIAs for luteinizing hormone and follicle-stimulating hormone (10), and Abraham developed the first E_2 RIA in 1969 (11). The RIA methodology developed by Yalow and Berson became highly successful because it offered a general system for measurement of an immensely wide range of compounds of clinical and biological importance.

The first steroid RIA method, which was developed by Abraham to quantify circulating levels of E_2 (11, 12), consisted of purification of E_2 in serum or plasma samples by organic solvent extraction and column chromatography, before its quantification by RIA. A small amount of tritiated E_2 was added to the samples first to correct for procedural losses. Organic solvent extraction removed the unconjugated steroids, leaving behind all water-soluble conjugated steroids (sulfates and glucuronides). Conjugated steroids are present in considerably higher concentrations than unconjugated steroids in blood. Thus, conjugated steroids such as estrone sulfate, dehydroepiandrosterone sulfate, and androsterone sulfate can cross-react with the antibody used in an RIA and cause overestimation in the measurement of a steroid hormone. Column chromatography was used to separate E_2 from E_1 and other

Table 1. Characteristics of E_2 assays used in large epidemiologic studies of bone health in postmenopausal women

Study	Assay method	E_2 detection limit (pg/mL)*	Lowest E_2 tertile or quartile (pg/mL)*
Cummings et al., 1998 (1) Ettinger et al., 1998 (2)	Extraction/chromatographic RIA†	5	<5 (quartile)
Garnero et al., 2000 (3)	Extraction/RIA‡	3	<11 (quartile)
Goderie-Plomp et al., 2004 (4)	Direct immunoassay§	1.3	<11 (tertile)
Rapuri et al., 2004 (5)	Direct immunoassay§	2.2	7.1 ± 0.14 (tertile)‖
Ettinger et al., 2004 (6)	Direct immunoassay¶	1.4	<2.7 (quartile)

*Multiply pg/mL by 3.67 to convert to pmol/L.
†Endocrine Sciences, Inc.
‡In-house method.
§Diagnostic Systems Laboratories.
‖Mean ± SD of lowest tertile.
¶Diagnostic Products Corporation.

unconjugated steroids, which include numerous E_2 metabolites. The E_2 RIA method was validated and shown to be sensitive, specific, precise, and accurate.

The extraction/chromatographic RIA method developed for E_2 was soon adapted to measure other steroid hormones. In addition, multiple analytes began to be measured in a relatively small sample volume. For example, androstenedione, DHEA, testosterone, E_2, and E_1 can be measured reliably in 0.6 mL of postmenopausal serum (13). In the last 35 years or so, the extraction/chromatographic RIA method has remained essentially the same and has been used in countless studies, yielding data that have enriched the field of endocrinology with new knowledge. In addition, the method has played an important role in measurements of a variety of steroid hormones used for diagnostic testing in clinical laboratories.

Direct Immunoassays. Although the extraction/chromatographic RIA method is highly reliable under most circumstances when thoroughly validated, it has some disadvantages. The method is cumbersome, time-consuming, and costly. For these reasons, direct RIAs, which do not incorporate a purification step before quantification of the analyte, were developed in the late 1970s. Kits that contained the necessary reagents to perform steroid RIAs became available commercially. Subsequently, the assay reagents were used in an instrument so that the assay could be done on an automated platform. This required replacement of the radioactive marker used in the assay with a nonradioactive one. Presently, direct immunoassay methods use chemiluminescent, fluorescent, or enzyme markers for quantification purposes.

Unlike extraction/chromatographic RIAs, which are time-consuming, direct immunoassay methods are simple, convenient, rapid, relatively inexpensive, and require less serum (usually only 0.1 mL). However, they can also have serious disadvantages. They often overestimate steroid concentrations due to lack of specificity of the antibody. For example, a major contributor to overestimation of levels in direct E_2 and E_1 immunoassays is estrone sulfate, which may cross-react substantially with E_2 and E_1 antisera used in immunoassays. This limitation is especially important in serum samples from postmenopausal women who have very low E_2 and E_1 levels but relatively high estrone sulfate levels. Also, matrix differences between the serum sample and solutions of standards may affect the results. This is particularly true of samples that are hemolyzed or lipemic. In addition, certain steroids (e.g., E_2 and testosterone) may not be released efficiently from sex hormone–binding globulin with the specific reagent used by the kit manufacturer to accomplish this. This could underestimate the levels. A particularly important and largely unrecognized problem is that direct assays lack sufficient sensitivity and specificity to measure low levels of steroids such as E_2, E_1, and testosterone in serum from postmenopausal women and both estrogens in male serum.

Evidence showing that certain direct steroid immunoassays carried out with reagents in commercial kits are not reliable can be found in several studies. In our study, we evaluated four different commercially available direct testosterone immunoassay kits and used our extraction/chromatographic testosterone RIA as a standard for comparison in 10 premenopausal and 10 postmenopausal women, as well as 10 men (14). Our results show that the four assays using the kits did generally well for male serum samples, but samples from premenopausal and postmenopausal women showed poor accuracy as they significantly deviated ($P < 0.05$) from the standard values based on ANOVA testing and/or showed poor intraclass correlations (assay correlations in premenopausal samples: 0.20, 0.52, 0.70, 0.71, and in postmenopausal samples: 0.62, 0.64, 0.88, 0.95). Our findings are consistent with those reported by Taieb et al. (15) who measured serum testosterone levels in women, men, and children by use of 10 different direct testosterone immunoassay kits and by isotope dilution gas chromatography-mass spectrometry (GC-MS). Compared with GC-MS, 7 of the 10 direct immunoassays overestimated testosterone concentrations in samples from women. The mean testosterone values measured by the immunoassays were 46% above those obtained by GC-MS. In samples from men, the direct immunoassays underestimated testosterone concentrations, giving mean results that were 12% below those obtained by GC-MS. On the basis of their results, they concluded that the direct assays were acceptable for measuring testosterone in male samples, but not in samples from women or children. In an accompanying editorial (16) on the article by Taieb et al., the editors concluded that "guessing seems to be nearly as good as most commercially available immunoassays and clearly superior to some."

In the same study in which we evaluated the reliability of direct testosterone immunoassay kits, we also compared female serum E_2 levels in 30 women determined by direct E_2 immunoassay kits to those obtained by our extraction/chromatographic E_2 RIA (14). Again, we found that the direct immunoassays were unreliable, as they deviated significantly ($P < 0.05$) from the standard values based on ANOVA testing and/or showed poor intraclass correlation (<0.85 and as low as 0.38). Our findings were substantiated by Nelson et al. (17), who showed that not only direct E_2 immunoassays, but also direct E_1 immunoassays were unreliable. For comparison, they used a liquid chromatography-tandem MS (LC-MS/MS) assay for simultaneous measurement of E_2 and E_1 in female plasma. At low E_2 concentrations, two of the three direct immunoassays showed poor agreement and the third direct immunoassay showed no agreement with LC-MS/MS. E_1 results obtained by direct RIA correlated moderately well with E_1 measured by LC-MS/MS (adjusted $r^2 = 0.67$). However, E_1 measurements obtained by an RIA method that used an organic solvent extraction step before the RIA showed good to excellent agreement with LC-MS/MS.

In a pilot study of elderly postmenopausal women with low E_2 levels (<20 pg/mL), we found that three extraction/chromatographic RIAs were highly correlated in their E_2 measurements, whereas four direct immunoassays were not (18). Table 2 shows a comparison of the lower detection limits and required serum volumes among these assays to highlight the variability in assay performance and feasibility for research studies. The extraction/chromatographic RIAs resembled more closely the values obtained with a comparison method of GC-MS/MS and one another (correlations ranging

Table 2. An example of differing characteristics among E_2 indirect and direct immunoassays

Assay method	E_2 detection limit (pg/mL)	Serum volume required (mL)	Correlation coefficient (95% CI) with GC-MS/MS*	Correlation coefficient (95% CI) with GC-MS/MS†
GC-MS/MS‡	0.6	1.0	—	—
Extraction/ chromatographic RIA§	2.0	1.0	0.88 (0.79-0.94)	—
Extraction only RIA‖	0.8	0.5	0.94 (0.88-0.97)	0.89 (0.81-0.93)
Extraction/ chromatographic RIA¶	1.8	0.5¶	0.91 (0.82-0.95)	—
Direct RIA**	1.4	0.2	0.83 (0.70-0.91)	0.65†† (0.52-0.76)
Direct RIA‡‡	2.2	0.2	0.70 (0.49-0.83)	—
Direct immunoassay§§	5.0	0.3	0.57 (0.31-0.75)	—
Direct immunoassay‖‖	2.7	0.07	0.71 (0.50-0.84)	—

*Reference 18 is a pilot study in 40 postmenopausal women.
† Reference 18 is a confirmatory study in 374 postmenopausal women.
‡ SFBC Taylor Technologies, Inc.
§Esoterix Endocrinology, Inc.
‖Royal Marsden Hospital Research Laboratory (M. Dowsett).
¶University of Southern California research laboratory (F.Z. Stanczyk); 0.8 mL serum used in ref. 19.
**Diagnostic Products Corporation.
†† Correlation significantly different from that between RIA and GC-MS/MS ($P < 0.01$).
‡‡ Diagnostic Systems Laboratories.
§§Roche Diagnostics, Inc.
‖‖Ortho-Clinical Diagnostics.

from 0.82 to 0.94) than did the direct immunoassay kits (correlations ranging from 0.38 to 0.83). A confirmatory study compared the direct immunoassay and the extraction-based RIA, which correlated best with GC-MS/MS in the pilot study, to GC-MS/MS in a larger sample of 374 older postmenopausal women not taking hormone therapy (18). The direct immunoassay correlated significantly weaker ($P < 0.01$) with GC-MS/MS than did the extraction-based RIA (Table 2). Another recent study compared very low E_2 concentrations using RIA with a preceding extraction step and two direct immunoassays in women who, as a result of taking an aromatase inhibitor, had very low circulating E_2 levels (19). At least 70% of the E_2 measured by the two direct assays were estimated to be an artifact. Furthermore, adding a preextraction step to the use of the direct assays led to measurements resembling those from the extraction-only RIA. In the latter study, it was concluded that the use of direct assays alone "is inappropriate and is likely to give aberrant guidelines."

In some instances, direct immunoassays may be preferred to RIAs with preceding purification steps. For example, measurement of endogenous levels of conjugated steroids, such as estrone sulfate, dehydroepiandrosterone sulfate, and 3α-androstanediol glucuronide by direct immunoassay can provide valuable data in breast cancer or prostate cancer studies, due to the relatively high concentrations of these compounds in blood. Use of direct immunoassays for measuring those hormones has a big advantage in providing data simply and rapidly. If a conventional RIA is used to quantify the same hormones, it is essential to split (by hydrolysis) the sulfate or glucuronide group off the molecule and then measure the unconjugated moiety. Such an assay is very cumbersome because it is essential first to remove the unconjugated steroid in the sample by organic solvent extraction, and then to hydrolyze the sulfate or glucuronide group off the molecule, using an enzyme (for estrone sulfate and 3α-androstanediol glucuronide) or acidic conditions (for dehydroepiandrosterone sulfate). After that, another organic solvent extraction step is required, followed by chromatography and quantification by RIA. Due to the multiple steps that are involved in such assays, procedural losses are often high. In addition, because the unconjugated steroid is being measured after hydrolysis of the conjugated steroid, a correction is often made to account for the difference in molecular weight between the conjugated and unconjugated steroid. Thus, in this case, a properly validated direct immunoassay is preferred over the conventional RIA. However, the best method for measuring a conjugated steroid is by a MS assay, because the entire molecule can be measured with high specificity and precision.

MS Assay Methods

In recent years, there has been a growing interest in standardization of steroid hormone measurements. It is generally accepted that to accomplish this, the best method to use is a GC-MS/MS or LC-MS/MS assay. GC-MS/MS and LC-MS/MS methods are powerful analytic techniques that combine the resolving power of GC or LC with the high sensitivity and specificity of the mass spectrometer. The mass spectrometer functions as a unique detector that provides structural information on individual analytes as they elute from the GC or LC column. The MS technique first involves ionization of the analyte at the ionization source. This is followed by separation and detection of the ions in the mass analyzer. A mass spectrum is produced in which the relative abundance of a particular ion is plotted as a function of the mass-to-charge ratio, and the concentration of the analyte is then obtained. MS/MS involves use of a collision cell in which the ion of interest (precursor ion) undergoes collision-induced fragmentation into product ions. The mass of the product ion is then determined at the detector.

There is a paucity of data comparing serum steroid hormone levels measured by RIA and MS assay methods. However, in the randomized clinical trial involving 374 postmenopausal women (18), we found that the extraction/chromatographic RIA [r = 0.50; 95% confidence interval (95% CI), 0.40-0.59] and GC-MS/MS (r = 0.52; 95% CI, 0.43-0.61) correlated better with body mass index (BMI), a biological indicator of estrogen, than the direct immunoassay (r = 0.29; 95% CI, 0.15-0.41; P = 0.01). The extraction/chromatographic RIA (r = 0.89; 95% CI, 0.81-0.93) correlated more closely with GC-MS/MS than did the direct immunoassay (r = 0.65; 95% CI, 0.52-0.76; $P < 0.01$). Consistent with these findings, the correlation between the direct immunoassay and the extraction/chromatographic RIA was 0.66 (95% CI, 0.54-0.75). The extraction/chromatographic RIA also correlated better on retest with GC-MS/MS than did the direct assay. Two commercial laboratories have shown that there is a good correlation between extraction/chromatographic RIAs and LC-MS/MS assays (20-22). However, steroid hormone levels obtained by GC-MS/MS or LC-MS/MS assays are generally lower than corresponding values achieved by RIAs, presumably due to the greater specificity of the former assays (20-23). New reference ranges will be required for steroid hormones measured by MS methods.

LC-MS/MS and GC-MS/MS methodologies have the capability for achieving not only high sensitivity and specificity, and measurement of multiple analytes in a small sample volume, but also high throughput of samples. MS assays are considered to become the gold standard for steroid hormone measurements. However, presently, the use of this method has been restricted to a relatively small number of laboratories. This is due to the high cost of the mass spectrometer, the requirement for a highly trained individual to operate the instrument, and the time required to develop assays for a variety of steroid hormones. Nevertheless, the initial cost of the instrumentation and assay development can be offset by the relatively rapid turnaround time of highly accurate steroid measurements, compared with the costly, cumbersome, and time-consuming extraction/chromatography RIA methodology.

Another very important advantage of MS assays is that, eventually, profiles of steroid hormones will be measured in a single aliquot of serum. Presently, this can be done with urinary estrogens. Recently, Xu et al. (24) developed a LC-MS/MS assay for measuring the absolute quantities of 15 endogenous estrogens in urine obtained from premenopausal or postmenopausal women and from men. The method requires a single hydrolysis/extraction/derivatization step before quantification by LC-MS/MS and uses only 0.5 mL of urine. Although this method is promising, it needs to be extended to serum for measurement of principal estrogens and androgens and their metabolites in women and men.

Sensitivity of E_2 Assay Methods

For measurements of E_2 in postmenopausal women, sensitivity is an important issue due to the extremely low serum E_2 levels found in a substantial number of these women. For example, in an important study on the association between low endogenous levels of serum E_2 and risk of hip and vertebral fractures among normal postmenopausal women, 81 of 247 subjects had undetectable E_2 levels, using an extraction/chromatographic RIA method with a level of sensitivity of 5 pg/mL (1). Accurate measurements of such low levels of E_2 might be important for prediction of risk of fractures (1), response to antiestrogens for prevention of breast cancer (25), and degree of E_2 suppression in women receiving aromatase inhibitor therapy (26).

To enhance assay sensitivity, *in vitro* recombinant DNA bioassays, with sensitivities ranging from 0.02 to 1 pg/mL, have been developed (27, 28). These assays detected 2- to 3-fold lower mean levels of E_2 than the extraction/chromatographic RIA method done on the same samples in normal postmenopausal women. For example, E_2 values were 1.95 ± 0.52 pg/mL by ultrasensitive yeast bioassay and 5.90 ± 1.40 pg/mL by extraction/chromatographic RIA in postmenopausal women with breast cancer (28). Also, the cell bioassay detected mean values of 3.23 ± 2.96 pg/mL versus 11.9 ± 12.0 by extraction-only/RIA in 30 normal postmenopausal women (28). However, the correlations between the E_2 values with the two methods were excellent (r = 0.84, r = 0.79). It is not known whether these bioassays more truly reflect circulating values or are subject to artifactual influences that lower the levels detected. Notwithstanding this issue, the ultrasensitive bioassays are too cumbersome for routine use and are relegated to research laboratories for special tasks such as assessing small increments of E_2 absorption from the vagina (29).

An important question then is whether extraction/chromatographic RIA or MS assay methods are more sensitive for measurement of E_2 in the lowest quartile samples. One GC-MS/MS assay has a level of sensitivity of 0.6 pg/mL and detected E_2 in all 30 samples from normal postmenopausal women and 45 samples from postmenopausal women with breast cancer (30). Our recent review examined seven different RIAs that had detection limits ranging from 0.8 to 5 pg/mL (Table 2; ref. 18). These observations suggest that MS methods might have superior operating levels of sensitivity for measurement of samples with low E_2 levels, where enhanced sensitivity is important. Further study is needed to directly compare the sensitivities of various extraction/chromatographic RIA methods with both GC-MS/MS and LC-MS/MS assay methods in this subgroup of postmenopausal women.

Standardization of Steroid Hormone Assays in Serum/Plasma

In the absence of an established gold standard against which to standardize steroid hormone assay methods, use of criterion-related validity testing, when possible, may be an additional practical approach to evaluate extraction/chromatographic RIAs, direct immunoassays, or MS assays. For example, BMI is a biological variable known to be significantly correlated with E_2 levels. We compared low E_2 measurements by different assay methods with BMI in our recent studies of elderly postmenopausal women (18). GC-MS/MS and extraction/chromatographic RIAs correlated best and comparably with BMI, whereas direct immunoassays correlated less well with BMI. The data suggest that perhaps a

correlation of >0.5 between BMI and low E_2 measurements by an assay method may be a useful criterion in evaluating and choosing an E_2 assay method. Further studies are needed to confirm this.

Investigators planning epidemiologic studies with steroid hormone measurements would benefit from conducting run-in pilot tests before committing large samples to a chosen assay. Such pilot tests would evaluate a selected assay or several possible assays against more than one reference criterion, such as MS and BMI, for accuracy and reliability. However, this is not sufficient for ultimate accuracy because variation often occurs over time in the course of a study, which may ultimately compromise study results.

As stated earlier, there seems to be general agreement that MS assays will become the gold standard for steroid hormone measurements. However, currently, there is a misconception that whenever current MS methods are used to quantify a steroid hormone, the resulting value is very accurate and is obtained by the gold standard. It is important to realize that interlaboratory differences exist in steroid hormone concentrations measured by MS assays. This is due to the fact that conditions used to perform the assay may vary from laboratory to laboratory. Assay conditions that can affect the measurements include derivative preparation, type of internal standard, calibrator purity, and type of instrumentation. Thus, to standardize MS assays for measurement of steroid hormones, it is essential that assay conditions (reagents, instruments, and procedure) be the same in all laboratories involved in standardizing an assay. In addition, it is essential to minimize intra-assay and intralaboratory variability. Assay standardization will require a great effort by participating laboratories.

Recommended Approach to Assay Standardization

We recommend establishment of standard pools of serum/plasma with levels of commonly measured steroid hormones to allow cross-comparison between assay methods on an international basis. The first step would be assessment of steroid levels in the pools by validated GC-MS/MS and LC-MS/MS assays in a centralized laboratory, overseen by a group similar to the College of American Pathologists, but much more stringent. This is especially important and urgent now that major clinical diagnostic laboratories, such as Quest Diagnostics' Nichols Institute and Esoterix, have converted most of their steroid hormone assays from extraction/chromatographic RIAs to LC-MS/MS assays. This change has affected the reference ranges for steroid hormones in those laboratories. Therefore, it will be essential to have standard serum/plasma steroid hormone pools available for cross-comparison of different assay methods. A centralized laboratory overseeing such a comparison would move the field of steroid hormone research forward. Investigators should express their support for such a laboratory at scientific meetings and in scientific journals.

For validation of individual assays, whether using direct immunoassay, extraction-only RIA, extraction/chromatographic RIA, or MS assay methods, it will be necessary to determine agreed upon confidence limits related to measurements of different steroid hormones in standard pools. This criterion would allow validation of sensitivity, specificity, precision, and accuracy of current steroid hormone assay methodology and would provide a surrogate for a true gold standard. The Endocrine Society in conjunction with the Center for Disease Control is currently exploring this concept and should be strongly encouraged to establish such a program (31).

References

1. Cummings SR, Browner WS, Bauer D, et al. Endogenous hormones and the risk of hip and vertebral fractures among older women. New Engl J Med 1998;339:733–8.
2. Ettinger B, Pressman A, Sklarin P, Bauer DC, Cauley JA, Cummings SR. Associations between low levels of serum estradiol, bone density, and fractures among elderly women: The Study of Osteoporotic Fractures. J Clin Endocrinol Metab 1998;83:2239–43.
3. Garnero P, Sornay-Rendu E, Claustrat B, Delmas PD. Biochemical markers of bone turnover, endogenous hormones, and the risk of fractures in postmenopausal women: The OFELY Study. J Bone Miner Res 2000;15:1526–36.
4. Goderie-Plomp HW, van der Klift M, de Ronde W, Hofman A, de Jong FH, Pols HA. Endogenous sex hormones, sex hormone-binding globulin, and the risk of incident vertebral fractures in elderly men and women: the Rotterdam Study. J Clin Endocrinol Metab 2004;89:3261–9.
5. Rapuri PB, Gallagher JC, Haynatzki G. Endogenous levels of serum estradiol and sex hormone binding globulin determine bone mineral density, bone remodeling, the rate of bone loss, and response to treatment with estrogen in elderly women. J Clin Endocrinol Metab 2004;89:4954–62.
6. Ettinger B, Genant HK, Steiger P, Madvig P. Low-dosage micronized 17 β-estradiol prevents bone loss in postmenopausal women. Am J Obstet Gynecol 1992;166:479–88.
7. Key TJ; Endogenous Hormones Breast Cancer Collaborative Group. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. J Natl Cancer Inst 2003;95:1218–26.
8. Yalow RS, Berson SA. Assay of plasma insulin in human subjects by immunological methods. Nature 1959;184:1648–9.
9. Yalow RS, Berson SA. Immunoassay of endogenous plasma insulin in man. J Clin Invest 1960;38:1157–75.
10. Odell WD, Ross GT, Rayford RL. RIA of luteinizing hormone in human plasma or serum: physiological studies. J Clin Invest 1967;46:248–55.
11. Abraham GE. Solid-phase radioimmunoassay of estradiol-17β. J Clin Endocrinol Metab 1969;29:866–70.
12. Abraham GE, Odell WD. Solid-phase radioimmunoassay of serum estradiol-17β: a semi-automated approach. In: Peron FG, Caldwell BV, editors. Immunologic methods in steroid determination. New York (NY): Appleton-Century-Crofts; 1970. p. 87–112.
13. Setiawan VW, Haiman CA, Stanczyk FZ, Marchand LL, Henderson BE. Racial/ethnic differences in postmenopausal endogenous hormones: the multiethnic cohort study. Cancer Epidemiol Biomarkers Prev 2006;15:1849–55.
14. Stanczyk FZ, Cho MM, Endres DB, Morrison JL, Patel S, Paulson RJ. Limitations of direct estradiol and testosterone immunoassay kits. Steroids 2003;68:1173–8.
15. Taieb J, Mathian B, Millot F, et al. Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography-mass spectrometry in sera from 116 men, women, and children. Clin Chem 2003;49:1381–95.
16. Herold DA, Fitzgerald RL. Immunoassays for testosterone in women: better than a guess? [Editorial]. Clin Chem 2003;49:1250–1.
17. Nelson RE, Grebe SK, O'Kane DJ, Singh R. Liquid chromatographic-tandem mass spectrometry assay for simultaneous measurement of estradiol and estrone in human plasma. Clin Chem 2004;50:373–84.
18. Lee JS, Ettinger B, Stanczyk FZ, et al. Comparison of methods to measure low serum estradiol levels in postmenopausal women. J Clin Endocrinol Metab 2006;91:3791–7.
19. Dowsett M, Folkerd E. Deficits in plasma oestradiol measurement in studies and management of breast cancer. Breast Cancer Res 2005;7:1–4.
20. Wagner A, Kalall T, Curtin W, Morr M, Chandler W, Grant R. LC-MS/MS analysis of steroids for clinical evaluation of endocrine disorders. Program of the 87th Annual Meeting of the Endocrine Society. San Diego (CA): 2005.
21. Goldman MM, Li S, Caulfield MP, Clarke NJ, Reitz RE. Detection and quantification of androstenedione, 17-hydroxyprogesterone, and progesterone in human/plasma by high turbulent flow liquid

chromatography—atmospheric pressure chemical ionization tandem mass spectrometry. Program of the 87th Annual Meeting of the Endocrine Society. San Diego (CA); 2005.
22. Goldman MM, Clarke NJ, Caulfield MP, Reitz RE. Direct quantification of total testosterone in human sera or plasma by high turbulent flow liquid chromatography-atmospheric pressure chemical ionization-tandem mass spectrometry (HTLC-APCI-MS/MS). Program of the 87th Annual Meeting of the Endocrine Society. San Diego (CA); 2005.
23. Hsing AW, Stanczyk FZ, Bélanger A, Schroeder P, Chang L, Falk RT, et al. Reproducibility of serum sex steroid assays in men by radioimmunoassay and mass spectrometry. Cancer Epidemiol Biomarker Prev 2007;16:1004–8.
24. Xu X, Veenstra TD, Fox SD, et al. Measuring fifteen endogenous estrogens simultaneously in human urine by high-performance liquid chromatography-mass spectrometry. Anal Chem 2005;77: 7757–65.
25. Cummings SR, Duong T, Kenyon E, Cauley JA, Whitehead M, Krueger KA. Serum estradiol level and risk of breast cancer during treatment with raloxifene. JAMA 2002;287:216–20.
26. Geisler J, Haynes B, Anker G, Dowsett M, Lonning PE. Influence of letrozole and anastrozole on total body aromatization and plasma estrogen levels in postmenopausal breast cancer patients evaluated in a randomized, cross-over study. J Clin Oncol 2002;20: 751–7.
27. Klein KO, Demers LM, Santner SJ, Baron J, Cutler GB, Jr., Santen RJ. Use of ultrasensitive recombinant cell bioassay to measure estrogen levels in women with breast cancer receiving the aromatase inhibitor, letrozole. J Clin Endocrinol Metab 1995;80:2658–60.
28. Wang S, Paris F, Sultan CS, et al. Recombinant cell ultrasensitive bioassay for measurement of estrogens in postmenopausal women. J Clin Endocrinol Metab 2005;90:1407–13.
29. Santen RJ, Pinkerton JV, Conaway M, et al. Treatment of urogenitial atrophy with low-dose estradiol: preliminary results. Menopause 2002;9:179–87.
30. Santen RJ, Demers L, Ohorodnik S, et al. Superiority of gas chromatography/tandem mass spectrometry assay (GC/MS/MS) for estradiol for monitoring of aromatase inhibitor therapy. Steroids 2007;72:666–71.
31. Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H. Utility, limitations, and pitfalls in measuring testosterone: An Endocrine Society position statement. J Clin Endocrinol Metab 2007;92:405–13.