Request # 17056139

MAY 05, 2005

Mail To:

Fordham Health Sciences Library (OhioLINK#547)

Interlibrary Loan

3640 Colonel Glenn Highway

Dayton, OH 45435-0001

ARIEL

DOCLINE: Journal Copy

Title:	Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer
Title Abbrev:	Cancer Epidemiol Biomarkers Prev
Citation:	2004 Jul;13(7):1257-60
Article:	Specimen allocation in longitudinal biomarker stud
Author:	Tworoger S; Yasui Y; Chang L; Stanczyk F; McTiernan A
NLM Unique ID:	9200608 Verify: PubMed
PubMed UI:	15247140
ISSN:	1055-9965 (Print) 1538-7755 (Electronic)
Publisher:	American Association for Cancer Research, Philadelphia, PA :
Copyright:	Copyright Compliance Guidelines
Authorization:	BARB
Need By:	N/A
Maximum Cost:	**$15.00**
Patron Name:	gLASER, rEBECCA - tn: 64616
Referral Reason:	Lacking
Library Groups:	RESOURCE
Phone:	1.937.775-4110
Fax:	1.937.775-2232
Email:	fill@www.libraries.wright.edu
Ariel:	130.108.121.58
Comments:	**GMR-RL PLEASE ARIEL IF POSSIBLE. THANKS**
Routing Reason:	Routed to OHUOSU in Serial Routing - cell 2
Received:	May 05, 2005 (05:38 PM EST)
Lender:	Ohio State University/ Columbus/ OH USA (OHUOSU)

Filled 5/5 for www

This material may be protected by copyright law (TITLE 17,U.S. CODE)

Bill to: OHUDAC

Fordham Health Sciences Library (OhioLINK#547)

Interlibrary Loan

3640 Colonel Glenn Highway

Dayton, OH 45435-0001

Short Communication

Specimen Allocation in Longitudinal Biomarker Studies: Controlling Subject-Specific Effects by Design

Shelley S. Tworoger,[1] Yutaka Yasui,[2] Lilly Chang,[3] Frank Z. Stanczyk,[3] and Anne McTiernan[2,4,5]

[1]Channing Laboratory, Harvard Medical School and Brigham and Women's Hospital, Boston, Massachusetts; [2]Cancer Prevention Research Program, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington; [3]Department of Obstetrics and Gynecology, Keck School of Medicine, University of Southern California, Los Angeles, California; and [4]Department of Epidemiology, School of Public Health and Community Medicine, and [5]Department of Medicine, School of Medicine, University of Washington, Seattle, Washington

Abstract

It is important to understand specimen allocation factors that may impact the validity and reliability of results in longitudinal studies examining within-person changes in biomarker levels. Using data from a randomized clinical trial of an exercise intervention in 136 postmenopausal women, we determined the effect of assaying the baseline and follow-up samples of some subjects in different batches on the intervention effect estimates for serum concentrations of estrone, estradiol, testosterone, androstenedione, and dehydroepiandrosterone. Twenty-five subjects had their baseline and 3-month follow-up samples and 50 subjects had their baseline and 12-month samples assayed in different batches; all other subjects had their baseline, 3-month, and 12-month samples assayed in the same batch. Subjects with split samples were reassayed with all samples in the same batch. We compared the estimated regression coefficient for the intervention effect using the split sample data with one estimated excluding the split sample data and one estimated replacing the split sample data with the reassayed data. The median percentage difference in the intervention effect estimate was 59.6% between using versus excluding the split sample data and 74.6% between using the split sample versus using the reassayed data. In general, the coefficients from the model including the split sample data were closer to zero and statistically less significant than those from the models excluding the split sample data or using the reassayed data. These results suggest that bias can be artificially introduced into intervention effect estimates of longitudinal studies if samples from a subject are not assayed in the same batch.   (Cancer Epidemiol Biomarkers Prev 2004;13(7):1257–60)

Introduction

Addressing issues of quality assurance (QA) when collecting biological specimens and conducting biomarker assays is important to ensure valid and reliable results. While many studies have focused on study designs that collect or analyze one sample per subject (1), less information exists about QA procedures for longitudinal studies in which the primary aim is to examine within-person changes. Because most longitudinal studies have many samples, the samples must be divided into multiple batches for assaying. This batching process can introduce an additional variability beyond the intrinsic, within-batch (intraassay) variation (i.e., batch-to-batch variation; refs. 2-6).

Ideally, batching will not add further variation beyond the intraassay variability; however, it is common for the batch-to-batch variability to be >0 (3, 4). If the samples for a specific participant are assayed in different batches, within-person comparisons could be biased because splitting subjects' samples introduces an additional variation to the within-person comparisons. Designing specimen allocation schemes to minimize the impact of this type of variability is important for obtaining valid and reliable results in longitudinal studies (2-4, 6).

The goal of this article is to better understand specimen allocation factors that may impact biomarker analyses in longitudinal studies that obtain multiple samples per subject. We used data from a randomized clinical trial investigating the effect of a yearlong moderate intensity exercise program versus control on serum levels of endogenous sex hormones in postmenopausal women (7). Among the main outcomes of the study are changes in serum estrone, estradiol, androstenedione, dehydroepiandrosterone (DHEA), and testosterone concentrations. We determined whether assaying samples from the same subject in different batches would influence the association between intervention/control group and changes in hormone concentrations.

Methods

Overview of the Study. The design of the study is described in detail elsewhere (7-9). Briefly, the study investigated the effect of a yearlong moderate intensity exercise intervention in postmenopausal women on serum levels of endogenous sex hormones and, secondarily, changes in weight, body mass index, percentage body fat, and immune function. We randomized 173 postmenopausal women, ages 50 to 75 years, who were sedentary (<20 minutes of exercise three times per week) and overweight (body mass index >25 or between 24.0 and 24.9 kg/m² and a percentage body fat >33%). Participants resided in the Seattle, WA metropolitan area.

The recruitment process identified potentially eligible women primarily via mass mailings and media advertisements (8). Interested women were screened for eligibility by a telephone interview. Major ineligibility criteria included using hormone replacement therapy, being too physically active, and having medical conditions contraindicating moderate-to-vigorous intensity exercise. Eligible women were scheduled for a screening clinic visit to collect baseline data. Those who successfully completed the screening process were randomized to an exercise intervention group (n = 87) or stretching control group (n = 86). Informed consent was obtained following the requirements of the Fred Hutchinson Cancer Research Center Institutional Review Board.

Subjects provided a 50 mL sample of blood after fasting for at least 12 hours, at the screening clinic visit before randomization, and at clinic visits 3 and 12 months postrandomization. Blood was processed within 1 hour of collection; serum, plasma, and buffy coats were aliquoted into 1.8 mL tubes and stored at −70°C in one freezer. Date and time of collection and time since last meal were recorded.

Laboratory Assays. All laboratory assays were performed at the Reproductive Endocrine Research Laboratory (University of Southern California) directed by one of the authors (F.Z. Stanczyk). We include results from the assays (n = 136 subjects) completed between March 2001 and February 2002; assays for the remaining subjects were completed at a later time. Samples were placed into batches such that, within each batch, the number of exercise and control subjects was approximately equal and the sample order was random; subjects were included in approximately chronological order of randomization to minimize bias by differing length of storage time. We created a serum pool from ineligible subjects who had provided a baseline blood sample; the samples used for this pool were from individuals who were postmenopausal and not taking hormone replacement therapy. Two specimens of the pooled sample were placed in each batch; hereafter, these will be called the pooled QA samples. These pooled QA samples were used to determine the assay coefficient of variation (CV). Laboratory personnel were blinded with regard to intervention/control status, which samples belonged to the same subject, and whether the sample was a pooled QA sample.

For 25 subjects (11 intervention and 14 control), the baseline sample was assayed in a different batch from the 3-month and 12-month samples, and in an additional 25 subjects (11 intervention and 14 control), the baseline

sample was assayed in a different batch from only the 12-month sample. Data from these subjects are hereafter called the split sample data. Otherwise, all samples from a subject were assayed in the same batch. In February 2002, subjects with samples split into multiple batches were reassayed such that all samples from a subject were assayed in the same batch; these data are hereafter called the reassayed data.

Androstenedione, DHEA, testosterone, estrone, and estradiol were quantified by sensitive and specific RIAs following organic solvent extraction and Celite column partition chromatography (10-13). Chromatographic separation of the five steroids was achieved by use of different concentrations of toluene in isooctane and ethyl acetate in isooctane. Intraassay and interassay CVs were determined using a random effects model to assess the variance components of the results from the pooled QA samples (2, 14) and approximate estimates derived by the delta method (15). The intraassay and interassay CVs were 12.4% and 17.6% for estrone, 12.4% and 15.8% for estradiol, 8.4% and 12.0% for testosterone, 6.1% and 11.6% for DHEA, and 7.4% and 9.8% for androstenedione.

Statistical Analysis. We assessed the effect of assaying the samples of some subjects in different batches on the estimates of the intervention effect on hormone outcomes. To do this, we compared the intervention effect estimate from models including the split sample data versus either excluding that data or using the reassayed data in which all subjects had their samples assayed in the same batch. We modeled the intervention effect at 3 and 12 months separately because a different number of subjects had their samples assayed in different batches at the two time points. Due to the longitudinal nature of the data, we used generalized estimating equations with a Gaussian error, identity link, and an unstructured working correlation matrix (16). We modeled the log-transformed hormone values with indicator variables for batch, month (baseline, 3 months, or 12 months), intervention group, and interactions between month and intervention group as covariates.

The magnitude of the effects of using the split sample data versus either excluding the split sample data or using the reassayed data was quantified by the percentage change in the interaction term regression coefficient, which is an estimate of the intervention effect, between the fits. To compare the model using versus excluding the split sample data, we calculated the difference in the regression coefficient of the interaction term between the two fits and divided by the regression coefficient obtained from the model excluding the split sample data:

$|\beta_{\text{excluding split sample data}} - \beta_{\text{including split sample data}}| / \beta_{\text{excluding split sample data}}$. The corresponding comparison for the model using the split sample data versus the reassayed data was $|\beta_{\text{using reassayed data}} - \beta_{\text{including split sample data}}| / \beta_{\text{using reassayed data}}$. We determined this for both the 3-month and 12-month comparisons.

Results

The regression coefficients of the intervention effect obtained when including the split sample data generally were closer to the null and statistically less significant

than those obtained either when excluding the split sample data or using the reassayed data (Table 1). The only exceptions were for the 12-month estradiol intervention effect estimate in which the coefficient moved away from zero when including versus excluding the split sample data and for the 12-month testosterone data in which the effect estimate was nearly zero when using the reassayed data but 0.0171 when using the split sample data. The intervention effect estimate crossed zero in two cases when comparing including versus excluding the split sample data and in two cases when comparing using the split sample versus reassayed data.

The regression coefficients changed appreciably when including the split sample data versus excluding those data or using the reassayed data even after adjusting for batch effects. The percentage difference between the intervention effect coefficients for the models including versus excluding the split sample data ranged from 31.8% for estradiol to 138.3% for testosterone (median 59.6%). Similarly, when using the reassayed data, the percentage difference ranged from 1.7% for androstenedione to 2,237% for testosterone (median 74.6%). The large percentage change for testosterone in both cases was due to the very small intervention effect estimates for that hormone. In general, the percentage differences were larger for the 12-month estimates than for the 3-month estimates. The largest absolute changes occurred for the largest intervention effect estimates.

Discussion

We were interested in elucidating specimen allocation factors that may impact biomarker analyses in longitudinal studies. The interassay CVs were larger than the intraassay CVs for all the assays, indicating that there is an added variability when using multiple assay batches. This suggests that additional measurement error could be introduced into the effect estimates of longitudinal studies that assess changes of these hormones within subjects. However, this type of measurement error would occur only if the samples from a subject were not assayed in the same batch.

Further, the estimate of the intervention effect changed appreciably when we included the data from subjects whose baseline and follow-up samples were assayed in different batches. Introducing this type of measurement error usually led to a conservative estimate of the intervention effects but sometimes changed the direction of the effect, although, in the latter situation, the effect estimates from one of the models was very close to the null. This is consistent with the literature, which suggests that measurement error in outcome measures of longitudinal studies can lead to bias (17, 18).

These data indicate that sample allocation is an extremely important design issue in longitudinal studies. Cohort or intervention studies primarily interested in assessing changes in biomarkers over time should assay the baseline and follow-up samples together; baseline samples should not be assayed alone to assess baseline cross-sectional associations more quickly. This concept may also be extended to other study designs. For example, in matched case-control studies, a case and its matched control or controls should be assayed in the same batch. A problem with batching may also occur in case-cohort designs in which a random sample of the cohort is assayed at one point and all cases are assayed at a later time. Bias may be introduced due to interassay variability or a systematic change in the overall mean levels, strengthening arguments to use nested case-control studies.

The subjects whose samples were assayed in different batches were chosen from the first 136 women randomized so that the assays could begin before the end of the trial. Slightly more control (n = 28) than intervention women (n = 22) had split samples, although the difference was not statistically significant. In addition, each comparison model (excluding the split sample data or

Table 1. Intervention effect estimates* (P value) at 3 and 12 months versus baseline, comparing models using the split sample data to models either excluding the split sample data or using the reassayed data[1]

Hormone	Subjects with Split Samples[‡]	Including Split Sample Data	Excluding Split Sample Data	Percentage Change[§]	Using Reassayed Data	Percentage Change[‖]
Month 3 vs baseline						
Estrone	25	−0.0164 (0.72)	−0.0354 (0.46)	53.7	−0.0764 (0.04)	78.5
Estradiol	25	−0.0703 (0.21)	−0.1031 (0.11)	31.8	−0.1155 (0.03)	39.1
Testosterone	25	0.0018 (0.96)	−0.0047 (0.89)	138.3	−0.0079 (0.78)	122.8
DHEA	24	0.0645 (0.31)	−0.1052 (0.14)	38.7	−0.0847 (0.16)	23.8
Androstenedione	25	−0.1068 (0.07)	−0.1603 (0.02)	33.4	−0.1087 (0.06)	1.7
Month 12 vs baseline						
Estrone	50	0.0022 (0.96)	−0.0286 (0.61)	107.7	0.0651 (0.10)	103.4
Estradiol	50	−0.0309 (0.48)	0.0191 (0.72)	61.8	−0.0416 (0.32)	25.7
Testosterone	50	0.0171 (0.64)	0.0401 (0.34)	57.4	−0.0008 (0.98)	2,237
DHEA	49	−0.0156 (0.81)	−0.0548 (0.49)	71.5	−0.0546 (0.39)	71.4
Androstenedione	50	−0.0029 (0.96)	−0.0142 (0.82)	79.6	−0.0131 (0.81)	77.9

*The intervention effect estimate is the interaction term comparing exercisers with controls for change in each hormone over time.
[†]Subjects whose samples originally were split into different batches were reassayed such that all samples from an individual subject were in the same batch.
[‡]The total number of subjects for estrone, estradiol, testosterone, and androstenedione is 136; the total number of subjects for DHEA is 135.
[§]Percentage change: |β_excluding split sample data − β_including split sample data| / β_excluding split sample data.
[‖]Percentage change: |β_using reassayed data − β_including split sample data| / β_using reassayed data.

using the reassayed data) has some limitations. Specifically, the model that excludes the split sample data has a different subset of subjects than the model including that data; this could partly explain the difference in the estimates rather than measurement error. Further, the model using the reassayed data has slightly different covariates than the model using the split sample data because of the additional indicator variables needed for the reassay batches. In this case, differences in the estimates could be partly due to the different covariate adjustment. However, despite these possible limitations, we noted a similar pattern of measurement error introduced by assaying the baseline and follow-up samples from a subject in different batches with both comparison models.

In conclusion, longitudinal studies using biomarker outcomes must be carefully designed to avoid unnecessary measurement error and variability. It is useful to include blinded replicates of a single, and preferably pooled, sample in each batch to determine the intraassay and interassay CVs. This information can be useful when interpreting the results of a trial, particularly if the expected changes are small or the effect estimates are not significant. In addition, all samples from an individual subject or matched set of subjects should be assayed in the same batch to reduce measurement error when estimating changes in outcomes over time and to better assure unbiased estimates.

References

1. Toniolo P, Boffetta P, Shuker DEG, Rothman N, Hulka B, Pearce N, editors. Application of biomarkers in cancer epidemiology. Lyon (France): IARC; 1997.
2. McShane LM, Dorgan JF, Greenhut S, Damato JJ. Reliability and validity of serum sex hormone measurements. Cancer Epidemiol Biomarkers & Prev 1996;5:923-8.
3. Gail MH, Fears TR, Hoover RN, et al. Reproducibility studies and interlaboratory concordance for assays of serum hormone levels:

estrone, estradiol, estrone sulfate, and progesterone. Cancer Epidemiol Biomarkers & Prev 1996;5:835-44.
4. Falk RT, Gail MH, Fears TR, et al. Reproducibility and validity of radioimmunoassays for urinary hormones and metabolites in pre- and postmenopausal women. Cancer Epidemiol Biomarkers & Prev 1999;8:567-77.
5. Bolelli G, Muti P, Micheli A, et al. Validity for epidemiological studies of long-term cryoconservation of steroid and protein hormones in serum and plasma. Cancer Epidemiol Biomarkers & Prev 1995;4:509-13.
6. Rinaldi S, Dechaud H, Biessy C, et al. Reliability and validity of commercially available, direct radioimmunoassays for measurement of blood androgens and estrogens in postmenopausal women. Cancer Epidemiol Biomarkers & Prev 2001;10:757-65.
7. McTiernan A, Ulrich CM, Yancey D, et al. The Physical Activity for Total Health (PATH) Study: rationale and design. Med Sci Sports Exerc 1999;31:1307-12.
8. Tworoger SS, Yasui Y, Ulrich CM, et al. Mailing strategies and recruitment into an intervention trial of the exercise effect on breast cancer biomarkers. Cancer Epidemiol Biomarkers & Prev 2002;11:73-7.
9. Irwin ML, Yasui Y, Ulrich CM, et al. Effect of exercise on total and intra-abdominal body fat in postmenopausal women: a randomized controlled trial. JAMA 2003;289:323-30.
10. Goebelsmann U, Bernstein GS, Gale JA, et al. Serum gonadotropin, testosterone, estradiol and estrone levels prior to and following bilateral vasectomy. In: Lepow IH, Crozier R, editors. Vasectomy: immunologic and pathophysiologic effects in animals and man. New York: Academic Press; 1979. p. 165.
11. Probst-Hensch NM, Ingles SA, Diep AT, et al. Aromatase and breast cancer susceptibility. Endocr Relat Cancer 1999;6:165-73.
12. Goebelsmann U, Arce JJ, Thorneycroft IH, Mishell DR Jr. Serum testosterone concentrations in women throughout the menstrual cycle and following HCG administration. Am J Obstet Gynecol 1974; 119:445-52.
13. Dorgan JF, Stanczyk FZ, Longcope C, et al. Relationship of serum dehydroepiandrosterone (DHEA), DHEA sulfate, and 5-androstene-3β, 17 β-diol to risk of breast cancer in postmenopausal women. Cancer Epidemiol Biomarkers & Prev 1997;6:177-81.
14. Rosner B. Fundamentals of biostatistics. Belmont (CA): Duxbury Press; 1995.
15. Chinn S. The assessment of methods of measurement. Stat Med 1990;9:351-62.
16. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986;42:121-30.
17. Yanez ND III, Kronmal RA, Shemanski LR, Psaty BM. A regression model for longitudinal change in the presence of measurement error. Ann Epidemiol 2002;12:34-8.
18. Armstrong BK, White E, Saracci R. Principles of exposure measurement in epidemiology. Monogr Epidemiol Biostat 1992;21:252-8.